

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 03:43:49 ; Search time 8.44444 Seconds
(without alignments)
16.652 Million cell updates/sec

Title: US-10-790-768A-1
Sequence: 1 RKMUKSTRQR 12

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_MA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	63.8	54	US-11-150-054A-13	Sequence 13, Appl
2	37	63.8	54	US-11-150-054A-21	Sequence 21, Appl
3	37	63.8	54	US-11-150-054A-26	Sequence 26, Appl
4	34	58.6	327	US-11-165-211-52	Sequence 52, Appl
5	34	58.6	327	US-11-165-226-62	Sequence 62, Appl
6	32	55.2	211	US-11-214-199-10	Sequence 10, Appl
7	32	55.2	212	US-11-214-199-4	Sequence 4, Appl
8	32	55.2	212	US-11-214-199-12	Sequence 12, Appl
9	32	55.2	411	US-11-092-168-8	Sequence 8, Appl
10	32	55.2	672	US-10-689-742-70	Sequence 70, Appl
11	31	53.4	27	US-11-078-469-32	Sequence 32, Appl
12	31	53.4	28	US-11-078-469-33	Sequence 33, Appl
13	31	53.4	30	US-11-078-469-24	Sequence 24, Appl
14	31	53.4	30	US-11-078-469-62	Sequence 62, Appl
15	31	53.4	119	US-11-120-308-146	Sequence 146, App
16	31	53.4	183	US-10-467-657-44	Sequence 44, App
17	31	53.4	183	US-10-467-657-8498	Sequence 8498, App
18	31	53.4	277	US-11-120-308-150	Sequence 150, App
19	31	53.4	280	US-10-821-234-1300	Sequence 1300, App
20	31	53.4	388	US-11-083-389-220	Sequence 220, App
21	31	53.4	407	US-11-051-267-25	Sequence 25, Appl
22	31	53.4	566	US-11-033-039-1244	Sequence 1244, App
23	31	53.4	580	US-11-054-281-95	Sequence 95, Appl
24	31	53.4	1029	US-10-821-234-908	Sequence 908, App
25	30	51.7	22	US-11-016-542-12	Sequence 12, Appl

26	30	51.7	46	US-10-467-657-3030	Sequence 3030, App
27	30	51.7	54	US-11-150-054A-3	Sequence 3, Appl
28	30	51.7	54	US-11-150-054A-44	Sequence 44, Appl
29	30	51.7	54	US-11-150-054A-53	Sequence 53, Appl
30	30	51.7	200	US-11-093-746A-19	Sequence 19, Appl
31	30	51.7	216	US-10-821-234-925	Sequence 925, App
32	30	51.7	255	US-10-467-657-5630	Sequence 5630, App
33	30	51.7	284	US-11-089-551A-36	Sequence 36, Appl
34	30	51.7	289	US-10-793-656-1120	Sequence 1120, App
35	30	51.7	441	US-11-024-959-410	Sequence 410, App
36	30	51.7	447	US-11-024-959-286	Sequence 286, App
37	30	51.7	510	US-11-024-959-409	Sequence 409, App
38	30	51.7	521	US-11-091-914-2	Sequence 2, Appl
39	30	51.7	884	US-10-995-561-786	Sequence 786, App
40	30	51.7	930	US-10-821-234-1188	Sequence 1188, App
41	29	50.0	12	US-10-877-961B-63	Sequence 63, Appl
42	29	50.0	21	US-10-895-064-2549	Sequence 2549, App
43	29	50.0	54	US-11-150-054A-32	Sequence 32, Appl
44	29	50.0	54	US-11-150-054A-35	Sequence 35, Appl
45	29	50.0	54	US-11-150-054A-48	Sequence 48, Appl

ALIGNMENTS

```
RESULT 1
US-11-150-054A-13
Sequence 13, Application US/11150054A
Publication No. US20050278801A1
GENERAL INFORMATION:
APPLICANT: Lasener, Michael
TITLE OF INVENTION: Placid Transit Peptides
FILE REFERENCE: 2119-4284US1
CURRENT FILING DATE: 2005-06-09
PRIOR APPLICATION NUMBER: US/11/150, 054A
PRIOR FILING DATE: 2004-06-09
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn version 3.3
SEQ ID NO 13
LENGTH: 54
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-11-150-054A-13

Query Match      63.8%; Score 37; DB 7; Length 54;
Best Local Similarity 58.3%; Pred. No. 1;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 RKMUKSTRQR 12
      |::|::|::|
DB      34 RFRNRTQR 45

RESULT 2
US-11-150-054A-21
Sequence 21, Application US/11150054A
Publication No. US20050278801A1
GENERAL INFORMATION:
APPLICANT: Lasener, Michael
APPLICANT: Wilkinson, Jack O.
TITLE OF INVENTION: Placid Transit Peptides
FILE REFERENCE: 2119-4284US1
CURRENT FILING DATE: 2005-06-09
PRIOR APPLICATION NUMBER: US/11/150, 054A
PRIOR FILING DATE: 2004-06-09
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn version 3.3
SEQ ID NO 21
```

LENGTH: 54
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-11-150-054A-21

Query Match 63.8%; Score 37; DB 7; Length 54;
Best Local Similarity 58.3%; Pred. No. 1;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQR 12
DB 34 RRFNRNTRRQR 45

RESULT 3
US-11-150-054A-26
Sequence 26, Application US/1150054A
Publication No. US20050278801A1
GENERAL INFORMATION:
APPLICANT: Lasener, Michael
APPLICANT: Wilkinson, Jack Q.
TITLE OF INVENTION: Plasticid Transit Peptides
FILE REFERENCE: 2119-4284US1
CURRENT APPLICATION NUMBER: US/11/150,054A
CURRENT FILING DATE: 2005-06-09
PRIOR APPLICATION NUMBER: 60/578,535
PRIOR FILING DATE: 2004-06-09
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn version 3.3
SEQ ID NO 26
LENGTH: 54
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-11-150-054A-26

Query Match 63.8%; Score 37; DB 7; Length 54;
Best Local Similarity 58.3%; Pred. No. 1;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQR 12
DB 34 RRFNRNTRRQR 45

RESULT 4
US-11-165-211-52
Sequence 52, Application US/11165211
Publication No. US20050287626A1
GENERAL INFORMATION:
APPLICANT: KYOMA HAKKO KOGYO CO., LTD.
TITLE OF INVENTION: Process for producing dipeptides
FILE REFERENCE: 4093-14
CURRENT APPLICATION NUMBER: US/11/165,211
CURRENT FILING DATE: 2005-06-24
PRIOR APPLICATION NUMBER: JP2004-189012
PRIOR FILING DATE: 2004-06-25
NUMBER OF SEQ ID NOS: 132
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 52
LENGTH: 327
TYPE: PRT
ORGANISM: Escherichia coli
US-11-165-211-52

Query Match 58.6%; Score 34; DB 7; Length 327;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KSTRRQR 11

DB 124 KSTRRQR 130

RESULT 5
US-11-165-226-62
Sequence 62, Application US/11165226
Publication No. US20050287627A1
GENERAL INFORMATION:
APPLICANT: KYOMA HAKKO KOGYO CO., LTD.
TITLE OF INVENTION: Process for producing dipeptides or dipeptide derivatives
FILE REFERENCE: 4093-13
CURRENT APPLICATION NUMBER: US/11/165,226
CURRENT FILING DATE: 2005-06-24
PRIOR APPLICATION NUMBER: JP2004-189007
PRIOR FILING DATE: 2004-06-25
NUMBER OF SEQ ID NOS: 131
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 62
LENGTH: 327
TYPE: PRT
ORGANISM: Escherichia coli
US-11-165-226-62

Query Match 58.6%; Score 34; DB 7; Length 327;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KSTRRQR 11
DB 124 KSTRRQR 130

RESULT 6
US-11-214-199-10
Sequence 10, Application US/11214199
Publication No. US2006003377A1
GENERAL INFORMATION:
APPLICANT: HILTON, Douglas J
APPLICANT: ALEXANDER, Warren S
APPLICANT: VINEY, Elizabeth M
APPLICANT: WILSON, Tracey A
APPLICANT: RICHARDSON, Rachael T
APPLICANT: STARR, Robyn
APPLICANT: NICHOLSON, Sandra E
APPLICANT: MERCALF, Donald
APPLICANT: NICOLA, Nicos A
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC PROTEINS COMPRISING A SOCS
FILE REFERENCE: 109762A
CURRENT APPLICATION NUMBER: US/11/214,199
CURRENT FILING DATE: 2005-08-29
PRIOR APPLICATION NUMBER: US/09/908,805
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 09/302,769
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 81
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 211
TYPE: PRT
ORGANISM: Human
US-11-214-199-10

Query Match 55.2%; Score 32; DB 7; Length 211;
Best Local Similarity 58.3%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQR 12
DB 159 RRMUGAPLRQR 170

```
RESULT 7
US-11-214-199-4
; Sequence 4, Application US/11214199
; Publication No. US2006000377A1
; GENERAL INFORMATION:
; APPLICANT: HILTON, Douglas J
; APPLICANT: ALEXANDER, Warren S
; APPLICANT: VINNEY, Elizabeth M
; APPLICANT: WILSON, Tracey A
; APPLICANT: RICHARDSON, Rachael T
; APPLICANT: STAR, Rodyn
; APPLICANT: NICHOLSON, Sandra E
; APPLICANT: METCALF, Donald
; APPLICANT: NICOLA, Nicos A
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC PROTEINS COMPRISING A SOCS
; FILE REFERENCE: BOX
; CURRENT FILING DATE: 2005-08-29
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US/09/908,805
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 09/302,769
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Mouse
US-11-214-199-4

Query Match      55.2%; Score 32; DB 7; Length 212;
Best Local Similarity 58.3%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 RKMKSTRRQR 12
      |||:|||||
Db      160 RRMGAPLRQR 171

RESULT 8
US-11-214-199-12
; Sequence 12, Application US/11214199
; Publication No. US2006000377A1
; GENERAL INFORMATION:
; APPLICANT: HILTON, Douglas J
; APPLICANT: ALEXANDER, Warren S
; APPLICANT: VINNEY, Elizabeth M
; APPLICANT: WILSON, Tracey A
; APPLICANT: RICHARDSON, Rachael T
; APPLICANT: STAR, Rodyn
; APPLICANT: NICHOLSON, Sandra E
; APPLICANT: METCALF, Donald
; APPLICANT: NICOLA, Nicos A
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC PROTEINS COMPRISING A SOCS
; FILE REFERENCE: BOX
; CURRENT FILING DATE: 2005-08-29
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US/09/908,805
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 09/302,769
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Rat
US-11-214-199-12

Query Match      55.2%; Score 32; DB 7; Length 212;
Best Local Similarity 58.3%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 RKMKSTRRQR 12
      |||:|||||
Db      160 RRMGAPLRQR 171

RESULT 9
US-11-092-168-8
; Sequence 8, Application US/11092168
; Publication No. US2005027658A1
; GENERAL INFORMATION:
; APPLICANT: Arizona Board of Regents on behalf of The University of Arizona
; APPLICANT: Montigen Pharmaceuticals, Inc.
; APPLICANT: Hurley, Laurence H.
; APPLICANT: Mahadevan, Daruka
; APPLICANT: Han, Haiyong
; APPLICANT: Bears, David J.
; APPLICANT: Vankayalapati, Hariprasad
; APPLICANT: Bashyam, Sridevi
; APPLICANT: Munoz, Ruben M.
; APPLICANT: Warner, Steven L.
; APPLICANT: Della Croce, Kimiko
; APPLICANT: Von Hoff, Daniel D.
; APPLICANT: Grand, Cory L.
; TITLE OF INVENTION: PROTEIN KINASE INHIBITORS
; FILE REFERENCE: 920214.00003CONT3
; CURRENT FILING DATE: 2005-03-29
; PRIOR FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: US 10/965,313
; US 60/608,529
; US 60/511,486
; US 60/511,485
; PRIOR FILING DATE: 2004-10-14
; 2004-09-09
; 2003-10-14
; 2003-10-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-092-168-8

Query Match      55.2%; Score 32; DB 7; Length 411;
Best Local Similarity 87.5%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 KMKSTR 9
      |||||
Db      43 KMKSTAR 50

RESULT 10
US-10-689-742-70
; Sequence 70, Application US/10689742
; Publication No. US20050250180A1
; GENERAL INFORMATION:
; APPLICANT: McCoy, John M
; APPLICANT: Lavalie, Edward R
; APPLICANT: Racie, Lisa A
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM
; FILE REFERENCE: 00766.000091.10
; CURRENT FILING DATE: 2003-10-22
; PRIOR FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US/10/689,742
; PRIOR FILING DATE: 2000-12-21
```

```
NUMBER OF SEQ ID NOS: 231
SOFTWARE: PatentIn version 3.2
SEQ ID NO 70
LENGTH: 672
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (107)..(107)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (111)..(111)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (117)..(118)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (645)..(645)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-689-742-70
```

```
Query Match          55.2%; Score 32; DB 6; Length 672;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 RMLKSTRRQR 12
|:::|:|:|
Db      598 RKLLEGRRESR 609
```

```
RESULT 11
US-11-078-469-32
; Sequence 32, Application US/11078469
; Publication No. US20050282755A1
; GENERAL INFORMATION:
; APPLICANT: HART, SCOTT A.
; APPLICANT: ZEH, KARIN
; APPLICANT: MACHLEIDT, THOMAS
; APPLICANT: STOLOW, DAVID
; TITLE OF INVENTION: COMPOSITIONS HAVING ANTIMICROBIAL ACTIVITY AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: ANS-2001-UT
; CURRENT APPLICATION NUMBER: US/11/078,469
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 60/554,526
; PRIOR FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: 60/618,948
; PRIOR FILING DATE: 2004-10-15
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 32
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide moiety
; NAME/KEY: MOD_RES
; LOCATION: (1)_RES
; OTHER INFORMATION: Ac-Thr
US-11-078-469-32
```

```
Query Match          53.4%; Score 31; DB 7; Length 27;
Best Local Similarity 41.7%; Pred. No. 5.9;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY      1 RMLKSTRRQR 12
|:::|:|:|
```

```
Db      16 RNFMRARRRR 27
```

```
RESULT 12
US-11-078-469-33
; Sequence 33, Application US/11078469
; Publication No. US20050282755A1
; GENERAL INFORMATION:
; APPLICANT: HART, SCOTT A.
; APPLICANT: ZEH, KARIN
; APPLICANT: MACHLEIDT, THOMAS
; APPLICANT: STOLOW, DAVID
; APPLICANT: CONGER, DEE
; TITLE OF INVENTION: COMPOSITIONS HAVING ANTIMICROBIAL ACTIVITY AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: ANS-2001-UT
; CURRENT APPLICATION NUMBER: US/11/078,469
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 60/554,526
; PRIOR FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: 60/618,948
; PRIOR FILING DATE: 2004-10-15
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 33
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide moiety
; NAME/KEY: MOD_RES
; LOCATION: (1)_RES
; OTHER INFORMATION: Ac-Thr
US-11-078-469-33
```

```
Query Match          53.4%; Score 31; DB 7; Length 28;
Best Local Similarity 41.7%; Pred. No. 6.1;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
OY      1 RMLKSTRRQR 12
|:::|:|:|
Db      16 RNFMRARRRR 27
```

```
RESULT 13
US-11-078-469-24
; Sequence 24, Application US/11078469
; Publication No. US20050282755A1
; GENERAL INFORMATION:
; APPLICANT: HART, SCOTT A.
; APPLICANT: ZEH, KARIN
; APPLICANT: MACHLEIDT, THOMAS
; APPLICANT: STOLOW, DAVID
; APPLICANT: CONGER, DEE
; TITLE OF INVENTION: COMPOSITIONS HAVING ANTIMICROBIAL ACTIVITY AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: ANS-2001-UT
; CURRENT APPLICATION NUMBER: US/11/078,469
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 60/554,526
; PRIOR FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: 60/618,948
; PRIOR FILING DATE: 2004-10-15
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 24
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
```

```
US-11-078-469-24
```

OTHER INFORMATION: peptide moiety
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Ac-Thr
US-11-078-469-24

Query Match 53.4%; Score 31; DB 7; Length 30;
Best Local Similarity 41.7%; Pred. No. 6.6;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQRR 12
Db 16 RNFWRARRRRR 27

RESULT 14
US-11-078-469-62
Sequence 62, Application US/11078469
Publication No. US20050282755A1
GENERAL INFORMATION:
APPLICANT: HART, SCOTT A.
APPLICANT: ZEH, KARIN
APPLICANT: MACLEIDT, THOMAS
APPLICANT: STOLOV, DAVID
APPLICANT: CONGER, DEE
TITLE OF INVENTION: COMPOSITIONS HAVING ANTIMICROBIAL ACTIVITY AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: ANS-2001-UT
CURRENT APPLICATION NUMBER: US/11/078,469
CURRENT FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: 60/554,526
PRIOR FILING DATE: 2004-03-18
PRIOR APPLICATION NUMBER: 60/618,948
PRIOR FILING DATE: 2004-10-15
NUMBER OF SEQ ID NOS: 294
SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 62
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide moiety
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Ac-Thr
US-11-078-469-62

Query Match 53.4%; Score 31; DB 7; Length 30;
Best Local Similarity 41.7%; Pred. No. 6.6;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQRR 12
Db 16 RNFWRARRRRR 27

RESULT 15
US-11-120-308-146
Sequence 146, Application US/11120308
Publication No. US20060005277A1
GENERAL INFORMATION:
APPLICANT: Famodu, Omolayo O.
APPLICANT: Forge, Charlie
APPLICANT: Miao, Guo-Hua
TITLE OF INVENTION: CDNA Encoding Polypeptides
FILE REFERENCE: BB-1365 US NA
CURRENT APPLICATION NUMBER: US/11/120,308
CURRENT FILING DATE: 2005-05-02
PRIOR APPLICATION NUMBER: US/10/078,770
PRIOR FILING DATE: 2002-02-19

PRIOR APPLICATION NUMBER: 09/614,188
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: 60/143,400
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/153,534
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: 60/161,223
PRIOR FILING DATE: 1999-10-22
PRIOR APPLICATION NUMBER: 60/159,878
PRIOR FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: 60/157,401
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: 60/143,419
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/143,409
PRIOR FILING DATE: 1999-07-12
NUMBER OF SEQ ID NOS: 196
SOFTWARE: Microsoft Office 97
SEQ ID NO 146
LENGTH: 119
TYPE: PRT
ORGANISM: Oryza sativa
US-11-120-308-146

Query Match 53.4%; Score 31; DB 7; Length 119;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 LKSTRRQRR 12
Db 41 MKRTRRDRR 49

Search completed: February 4, 2006, 03:47:58
Job time : 9.44444 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 03:26:33 ; Search time 248.444 Seconds
(without alignments)
21.222 Million cell updates/sec

Title: US-10-790-768A-1
Perfect score: 58
Sequence: 1 RKMLKSTRQR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

A_Geneseq_21:*

- 1: geneseqp1980s:*
- 2: geneseqp1980s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	12	8	ADSL17617 Amino aci
2	58	100.0	15	8	ADSL17632 Amino aci
3	58	100.0	15	8	ADSL17630 Amino aci
4	58	100.0	15	8	ADSL17618 Amino aci
5	58	100.0	21	8	ADSL17636 Amino aci
6	58	100.0	21	8	ADSL17620 Amino aci
7	58	100.0	21	8	ADSL17635 Amino aci
8	58	100.0	27	8	ADSL17621 Amino aci
9	48	82.8	19	8	ADSL17634 Amino aci
10	48	82.8	25	8	ADSL17638 Amino aci
11	48	82.8	25	8	ADSL17637 Amino aci
12	43	74.1	63	4	AUS05025 Propionib
13	43	74.1	63	4	ABM47044 Propionib
14	43	74.1	247	7	ADJ70413 Human hea
15	42	72.4	33	4	AAG90210 C glutami
16	39	67.2	33	5	ABU05628 M. tuberc
17	39	67.2	33	5	ABU05754 M. tuberc
18	39	67.2	84	4	AU63875 Propionib
19	39	67.2	84	6	ABM60394 Propionib
20	39	67.2	93	4	AU63362 Propionib
21	39	67.2	93	4	AU67333 Propionib
22	39	67.2	93	6	ABM63852 Propionib
23	39	67.2	93	6	ABM59881 Propionib
24	39	67.2	131	3	AAG04024 Human sec

25	39	67.2	258	8	ADN22807 Bacterial
26	39	67.2	333	8	ADR66454 Aspergill
27	38	65.5	50	8	ADQ39534 Human myo
28	37	63.8	229	4	ABB16808 Human ner
29	37	63.8	402	4	AAM39344 Human poi
30	37	63.8	402	8	ADO79175 Ado79175 KLF12 ful
31	37	63.8	402	8	ADQ96310 T cell ac
32	37	63.8	402	8	ADP54643 Human PRO
33	37	63.8	402	8	ADQ7429 Cyclin-de
34	37	63.8	651	5	ABB93099 Herbicida
35	36	62.1	151	4	ABB68911 Drosophil
36	36	62.1	161	7	ABO78893 Pseudomon
37	36	62.1	250	6	ABU18327 Protein e
38	36	62.1	537	7	ABO77212 Pseudomon
39	36	62.1	708	7	ADJ69966 Human hea
40	35	60.3	81	5	ABP05731 Human ORF
41	35	60.3	157	2	AAW33886 Flea gall
42	35	60.3	157	2	AAW82360 Flea gall
43	35	60.3	220	7	ADH85947 Enterococ
44	35	60.3	287	8	ADQ67524 Novel hum
45	35	60.3	310	3	AAG27360 Arabidops

ALIGNMENTS

RESULT 1
ADSL17617
ID ADSL17617 standard; peptide; 12 AA.
AC ADSL17617;
DT 02-DEC-2004 (first entry)
DE Amino acid sequence of protein transduction domain (PTD) peptide #1.
XX protein transduction domain; PTD; PTD-cargo moiety complex;
XX cell immortalisation; cell viability; internalising peptide;
XX cell transport.
XX
XX
XX Synthetic.
XX W02004078933-A2.
XX 16-SEP-2004.
XX 04-MAR-2004; 2004WO-US006445.
XX 04-MAR-2003; 2003US-0451243P.
XX 03-MAR-2004; 2004US-00790768.
XX (BIOV-) BIONEER TECHNOLOGIES INC.
XX Karas M;
XX WPI; 2004-653708/63.
XX New isolated and purified polypeptide with a protein transduction domain,
XX useful for delivering small molecules, proteins and nucleic acids to an
XX intracellular compartment of a cell.
XX
XX Claim 1; SEQ ID NO 1; 60pp; English.
XX The present sequence represents a peptide which functions as a protein
XX transduction domain (PTD), and is capable of delivering small molecules,
XX proteins, and nucleic acids to an intracellular compartment of a cell. An
XX amino terminal lysine linker improves the efficiency of the PTD. The PTD
XX can be used in PTD-cargo moiety complexes that can reversibly immortalise
XX cells and increase cell viability in culture. The present sequence is a
XX reverse isomer of ADSL1761, a previously identified internalising
XX peptide.
XX Sequence 12 AA;
SQ

Query Match 100.0%; Score 58; DB 8; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.0063;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQRR 12
 |||||
 DB 1 RKMLKSTRRQRR 12

RESULT 2

ADSI17632
 ID ADSI17632 standard; peptide; 15 AA.

AC ADSI17632;

DT 02-DEC-2004 (first entry)

DE Amino acid sequence of a PTD designated Biotin-InvrP5.

KM protein transduction domain; PTD; PTD-cargo moiety complex;
 KW cell immortalisation; cell viability; internalising peptide;
 KW cell transport.

OS Synthetic.

PH Key Location/Qualifiers

FT Modified-site 1 /note= "biotin attached"

XX MO2004078933-A2.

XX 16-SEP-2004.

XX 04-MAR-2004; 2004MO-US006445.

XX 04-MAR-2003; 2003US-0451243P.

XX 03-MAR-2004; 2004US-00790768.

XX (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.

XX Karas M;

XX WPI; 2004-653708/63.

PT New isolated and purified polypeptide with a protein transduction domain,
 PT useful for delivering small molecules, proteins and nucleic acids to an
 PT intracellular compartment of a cell.

XX Example 4; SEQ ID NO 16; 60pp; English.

CC The specification describes peptides which function as a protein
 CC transduction domain (PTD), and are capable of delivering small molecules,
 CC proteins, and nucleic acids to an intracellular compartment of a cell. An
 CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
 CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
 CC cells and increase cell viability in culture. The present sequence
 CC represents PTD peptide used in an experiment to determine the
 CC transduction efficiency of inverted isomers of PTDs.

XX Sequence 15 AA;

Query Match 100.0%; Score 58; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0078;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQRR 12
 |||||
 DB 4 RKMLKSTRRQRR 15

RESULT 3

ADSI17630

ID ADSI17630 standard; peptide; 15 AA.

XX ADSI17630;

DT 02-DEC-2004 (first entry)

DE Amino acid sequence of a PTD designated Biotin-Lys-InvrP5.

KM protein transduction domain; PTD; PTD-cargo moiety complex;
 KW cell immortalisation; cell viability; internalising peptide;
 KW cell transport.

OS Synthetic.

PH Key Location/Qualifiers

FT Modified-site 1 /note= "biotin attached"

XX MO2004078933-A2.

XX 16-SEP-2004.

XX 04-MAR-2004; 2004MO-US006445.

XX 04-MAR-2003; 2003US-0451243P.

XX 03-MAR-2004; 2004US-00790768.

XX (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.

XX Karas M;

XX WPI; 2004-653708/63.

PT New isolated and purified polypeptide with a protein transduction domain,
 PT useful for delivering small molecules, proteins and nucleic acids to an
 PT intracellular compartment of a cell.

XX Example 4; SEQ ID NO 14; 60pp; English.

CC The specification describes peptides which function as a protein
 CC transduction domain (PTD), and are capable of delivering small molecules,
 CC proteins, and nucleic acids to an intracellular compartment of a cell. An
 CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
 CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
 CC cells and increase cell viability in culture. The present sequence
 CC represents PTD peptide used in an experiment to determine the
 CC transduction efficiency of inverted isomers of PTDs.

XX Sequence 15 AA;

Query Match 100.0%; Score 58; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0078;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQRR 12
 |||||
 DB 4 RKMLKSTRRQRR 15

RESULT 4

ADSI17618
 ID ADSI17618 standard; peptide; 15 AA.

AC ADSI17618;

DT 02-DEC-2004 (first entry)

DE Amino acid sequence of protein transduction domain (PTD) peptide #2.

KM protein transduction domain; PTD; PTD-cargo moiety complex;
 KW cell immortalisation; cell viability; internalising peptide;
 KW cell transport.

OS Synthetic.
XX WO2004078933-A2.
PN
XX
PD 16-SEP-2004.
XX
PF 04-MAR-2004; 2004WO-US006445.
XX
PR 04-MAR-2003; 2003US-0451243P.
XX 03-MAR-2004; 2004US-00790768.
PA (BIOV-) BIONEER TECHNOLOGIES INC.
XX
PI Karas M;
XX
DR WPI; 2004-653708/63.
XX
XX New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
PS Claim 3; SEQ ID NO 2; 60pp; English.
XX
XX The present sequence represents a peptide which functions as a protein
CC transduction domain (PTD), and is capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. The present PTD has a
CC lysine linker.
XX
SQ Sequence 15 AA;
XX
OY 1 RKMLKSTRRQRR 12
DB 4 RKMLKSTRRQRR 15
XX
RESULT 5
ADSI7636
ID ADSI7636 standard; peptide; 21 AA.
XX
AC ADSI7636;
XX
DT 02-DEC-2004 (first entry)
XX
DE Amino acid sequence of a PTD designated peptide 15.
XX
XX protein transduction domain; PTD; PTD-cargo moiety complex;
KW cell immortalisation; cell viability; internalising peptide;
KW cell transport.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
PT Modified-site 1
PT /note= "biotin attached"
XX
PN WO2004078933-A2.
XX
PD 16-SEP-2004.
XX
PF 04-MAR-2004; 2004WO-US006445.
XX
PR 04-MAR-2003; 2003US-0451243P.
XX 03-MAR-2004; 2004US-00790768.
PA (BIOV-) BIONEER TECHNOLOGIES INC.
XX

PI Karas M;
XX
XX WPI; 2004-653708/63.
DR
XX
XX New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
PS Example 10; SEQ ID NO 20; 60pp; English.
XX
XX The specification describes peptides which function as a protein
CC transduction domain (PTD), and are capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. ADSI7634-ADSI7638 represent
CC PTD peptides used in an experiment to determine if introduction of a
CC nuclear localisation signal affects translocation of a PTD-cargo moiety
CC complex.
XX
SQ Sequence 21 AA;
XX
OY 1 RKMLKSTRRQRR 12
DB 10 RKMLKSTRRQRR 21
XX
RESULT 6
ADSI7620
ID ADSI7620 standard; peptide; 21 AA.
XX
AC ADSI7620;
XX
DT 02-DEC-2004 (first entry)
XX
DE Amino acid sequence of protein transduction domain (PTD) peptide #3.
XX
XX protein transduction domain; PTD; PTD-cargo moiety complex;
KW cell immortalisation; cell viability; internalising peptide;
KW cell transport.
XX
OS Synthetic.
XX
PN WO2004078933-A2.
XX
PD 16-SEP-2004.
XX
PF 04-MAR-2004; 2004WO-US006445.
XX
PR 04-MAR-2003; 2003US-0451243P.
XX 03-MAR-2004; 2004US-00790768.
PA (BIOV-) BIONEER TECHNOLOGIES INC.
XX
PI Karas M;
XX
DR WPI; 2004-653708/63.
XX
XX New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
PS Disclosure; SEQ ID NO 4; 60pp; English.
XX
XX The present sequence represents a peptide which functions as a protein
CC transduction domain (PTD), and is capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise

CC cells and increase cell viability in culture. The present PTD has a
CC lysine linker and a nuclear localisation signal.

XX Sequence 21 AA;
SQ

Query Match 100.0%; Score 58; DB 8; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQRR 12
|||||
DB 4 RKMLKSTRRQRR 15

RESULT 7
ADS17635
ID ADS17635 standard; peptide; 21 AA.
XX
AC ADS17635;
XX

DT 02-DEC-2004 (first entry)
XX

DE Amino acid sequence of a PTD designated peptide 14.

XX protein transduction domain; PTD; PTD-cargo moiety complex;
KM cell immortalisation; cell viability; internalising peptide;
KW cell transport.
XX

OS Synthetic.
XX

FM Key Location/Qualifiers
PT Modified-site 1 /note= "biotin attached"

XX MO2004078933-A2.
PN
PD 16-SEP-2004.
XX

PF 04-MAR-2004; 2004WO-US006445.
XX

PR 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.
XX

PA (BIOV-) BIONEERING TECHNOLOGIES INC.
XX

PI Karas M;
XX

DR WPI; 2004-653708/63.
XX

PT New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.

XX Example 10; SEQ ID NO 19; 60pp; English.
PS

XX The specification describes peptides which function as a protein
CC transduction domain (PTD), and are capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. ADS17634-ADS17638 represent
CC PTD peptides used in an experiment to determine if introduction of a
CC nuclear localisation signal affects translocation of a PTD-cargo moiety
CC complex.
CC

XX Sequence 21 AA;
SQ

Query Match 100.0%; Score 58; DB 8; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQRR 12
|||||

DB 4 RKMLKSTRRQRR 15

RESULT 8
ADS17621
ID ADS17621 standard; peptide; 27 AA.
XX
AC ADS17621;
XX

DT 02-DEC-2004 (first entry)
XX

DE Amino acid sequence of protein transduction domain (PTD) peptide #4.

XX protein transduction domain; PTD; PTD-cargo moiety complex;
KM cell immortalisation; cell viability; internalising peptide;
KW cell transport.
XX

OS Synthetic.
XX

PN MO2004078933-A2.
XX

PD 16-SEP-2004.
XX

PF 04-MAR-2004; 2004WO-US006445.
XX

PR 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.
XX

PA (BIOV-) BIONEERING TECHNOLOGIES INC.
XX

PI Karas M;
XX

DR WPI; 2004-653708/63.
XX

PT New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.

XX Disclosure; SEQ ID NO 5; 60pp; English.
PS

XX The present sequence represents a peptide which functions as a protein
CC transduction domain (PTD), and is capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. The present PTD has a
CC lysine linker and 2 nuclear localisation signals.
CC

XX Sequence 27 AA;
SQ

Query Match 100.0%; Score 58; DB 8; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQRR 12
|||||
DB 10 RKMLKSTRRQRR 21

RESULT 9
ADS17634
ID ADS17634 standard; peptide; 19 AA.
XX
AC ADS17634;
XX

DT 02-DEC-2004 (first entry)
XX

DE Amino acid sequence of a PTD designated peptide 13.

XX protein transduction domain; PTD; PTD-cargo moiety complex;
KM cell immortalisation; cell viability; internalising peptide;
KW cell transport.
XX

OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 1 /note= "biotin attached"
XX	
PN	WO2004078933-A2.
XX	
PD	16-SEP-2004.
XX	
PE	04-MAR-2004; 2004WO-US006445.
XX	
PR	04-MAR-2003; 2003US-0451243P.
PR	03-MAR-2004; 2004US-00790768.
XX	
PA	(BIOW-) BIOWITTAKER TECHNOLOGIES INC.
XX	
P1	Karas M,
XX	
DR	WPI; 2004-653708/63.
XX	
PT	New isolated and purified polypeptide with a protein transduction domain, useful for delivering small molecules, proteins and nucleic acids to an intracellular compartment of a cell.
PT	
PS	Example 10; SEQ ID NO 18; 60pp; English.
XX	
CC	The specification describes peptides which function as a protein transduction domain (PTD), and are capable of delivering small molecules, proteins, and nucleic acids to an intracellular compartment of a cell. An amino terminal lysine linker improves the efficiency of the PTD. The PTD can be used in PTD-cargo moiety complexes that can reversibly immortalise cells and increase cell viability in culture. ADS17634-ADS17638 represent PTD peptides used in an experiment to determine if introduction of a nuclear localisation signal affects translocation of a PTD-cargo moiety complex.
CC	
CC	
CC	
SO	Sequence 19 AA;
XX	
OY	Query Match 82.8%; Score 48; DB 8; Length 19; Best Local Similarity 100.0%; Pred. No. 0.45; Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0
DB	3 MLKSTRRQRR 12 10 MLKSTRRQRR 19
RESULT 10	
ADSI17638	
ID	ADSI17638 standard; peptide; 25 AA.
XX	
AC	ADSI17638;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	Amino acid sequence of a PTD designated peptide 17.
XX	
KM	protein transduction domain; PTD; PTD-cargo moiety complex; cell immortalisation; cell viability; internalising peptide;
KW	cell transport.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 1 /note= "biotin attached"
XX	
PN	WO2004078933-A2.
XX	
PD	16-SEP-2004.
XX	
PF	04-MAR-2004; 2004WO-US006445.

XX	04-MAR-2003; 2003US-0451243P.
PR	03-MAR-2004; 2004US-00790768.
PA	(BIOW-) BIOWHITTAKER TECHNOLOGIES INC.
PI	Karas M;
XX	WPI; 2004-653708/63.
DR	
XX	
PT	New isolated and purified polypeptide with a protein transduction domain, useful for delivering small molecules, proteins and nucleic acids to an intracellular compartment of a cell.
PT	
XX	
PS	Example 10; SEQ ID NO 22; 60pp; English.
XX	
CC	The specification describes peptides which function as a protein transduction domain (PTD), and are capable of delivering small molecules, proteins, and nucleic acids to an intracellular compartment of a cell. An amino terminal lysine linker improves the efficiency of the PTD. The PTD can be used in PTD-cargo moiety complexes that can reversibly immortalise cells and increase cell viability in culture. ADS17634-ADS17638 represent PTD peptides used in an experiment to determine if introduction of a nuclear localisation signal affects translocation of a PTD-cargo moiety complex.
CC	
XX	
SQ	Sequence 25 AA;
Query Match	82.8%; Score 48; DB 8; Length 25;
Best Local Similarity	100.0%; Pred. No. 0.58;
Matches	10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	3 MLKSTRRQRR 12
DB	16 MLKSTRRQRR 25
RESULT 11	
ADS17637	
ID	ADS17637 standard; peptide; 25 AA.
XX	
AC	ADS17637;
XX	
DT	02-DEC-2004 (first entry)
DE	
XX	Amino acid sequence of a PTD designated peptide 16.
XX	
KW	protein transduction domain; PTD, PTD-cargo moiety complex; cell immortalisation; cell viability; internalising peptide; cell transport.
XX	
OS	Synthetic.
XX	
TH	Key
FT	Modified-site 1 Location/Qualifiers
FT	/note= "biotin attached"
XX	
PN	WO2004078933-A2.
XX	
PD	16-SEP-2004.
XX	
PF	04-MAR-2004; 2004WO-US006445.
PR	04-MAR-2003; 2003US-0451243P.
PR	03-MAR-2004; 2004US-00790768.
XX	
PA	(BIOW-) BIOWHITTAKER TECHNOLOGIES INC.
XX	
PI	Karas M;
DR	WPI; 2004-653708/63.
XX	
PT	New isolated and purified polypeptide with a protein transduction domain,

PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
PS Example 10; SEQ ID NO 21; 60pp; English.
XX
CC The specification describes peptides which function as a protein
CC transduction domain (PTD), and are capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. ADS17634-ADS17638 represent
CC PTD peptides used in an experiment to determine if introduction of a
CC nuclear localisation signal affects translocation of a PTD-cargo moiety
CC complex.
XX
SQ Sequence 25 AA;
XX
Query Match 82.8%; Score 48; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 3 MLKSTRRQR 12
Db 10 MLKSTRRQR 19
XX
RESULT 12
AAU50525
ID AAU50525 standard; protein; 63 AA.
XX
AC AAU50525;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #11421.
XX
KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59549.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 11720; 1069pp; English.
XX
CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 63 AA;
XX
Query Match 74.1%; Score 43; DB 4; Length 63;
Best Local Similarity 75.0%; Pred. No. 9.4;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 RKMLKSTRRQR 12
Db 48 RKMLKSTRRQR 59
XX
RESULT 13
ABM47044
ID ABM47044 standard; protein; 63 AA.
XX
AC ABM47044;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #11720.
XX
KM Acne vulgaris; anti-seborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Valilleve-Douglas J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64478.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 11720; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM5624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared

CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 63 AA;

SO
Query Match 74.1%; Score 43; DB 6; Length 63;
Best Local Similarity 75.0%; Pred. No. 9.4;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RKMLKSTRRORR 12
Db 48 RKMLKSTRRORR 59

RESULT 14
ADJ70413
ID ADJ70413 standard; protein: 247 AA.

XX ADJ70413;

DT 06-MAY-2004 (first entry)

XX Human heat mitochondrial protein as a therapeutic target SegID2219.

XX mitochondria; human; screening assay; diabetes mellitus;

KM Huntington's disease; osteoarthritis;

KM Leber's hereditary optic neuropathy; LHON;

KM mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

KM myoclonic epilepsy, ragged red fibre syndrome; MERRF; cancer;

KM neuroprotective; neurotrophic; antidiabetic; anticonvulsant; antiarthritic;

KM osteoprotective; ophthalmological; cytostatic.

XX Homo sapiens.

XX NO2003087768-A2.

XX 23-OCT-2003.

XX 04-APR-2003; 2003WO-US010670.

XX 12-APR-2002; 2002US-0372843P.

XX 17-JUN-2002; 2002US-0389987P.

XX 20-SEP-2002; 2002US-0412418P.

XX (MITO-) MITOKOR.

XX (BUCK-) BUCK INST AGE RES.

XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM,

XX Warnock DE;

XX WPI; 2003-845369/78.

XX Identifying a mitochondrial target for drug screening assays and for
XX treating diseases associated with altered mitochondrial function,
XX PT comprises detecting a modified polypeptide in a sample and correlating
XX with the disease.

PS Claim 1; SEQ ID NO 2219; 180pp; English.

XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, neurotrophic, antidiabetic,
CC anticonvulsant, antiarthritic, osteoprotective, ophthalmological and
CC cyostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.

XX Sequence 247 AA;

SO
Query Match 74.1%; Score 43; DB 7; Length 247;
Best Local Similarity 81.8%; Pred. No. 34;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RKMLKSTRRORR 11
Db 130 RKSLKSTRRRR 140

RESULT 15
AAG90210
ID AAG90210 standard; protein: 33 AA.

XX AAG90210;

DT 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 3964.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KM organic acid synthesis.

XX Corynebacterium glutamicum.

XX BP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H,

XX Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

XX N-PSDB; AAH65429.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
CC mutation point of a gene, measuring expression of a gene, analyzing
CC expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 3964; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from

CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX

SQ Sequence 33 AA;

Query Match 72.4%; Score 42; DB 4; Length 33;

Best Local Similarity 75.0%; Pred. No. 7.5;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRORR 12

DB 18 RKMLRRTTRVQRR 29

Search completed: February 4, 2006, 03:36:12
 Job time : 253.944 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 03:36:41 ; Search time 22.6667 Seconds
(without alignments)
50.938 Million cell updates/sec

Title: US-10-790-768a-1
Perfect score: 58
Sequence: 1 RKMKSTRRQRR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	72.4	32	2	T36275
2	33	67.2	33	2	A87213
3	39	67.2	258	2	T24499
4	38	65.5	361	2	A81979
5	37	63.8	651	2	T10219
6	36	62.1	69	2	S39424
7	36	62.1	89	2	T29995
8	35	60.3	110	2	S15153
9	35	60.3	308	2	A85904
10	35	60.3	308	2	H65035
11	35	60.3	308	2	C91059
12	35	60.3	829	2	A40894
13	35	60.3	1017	2	T31354
14	35	60.3	1020	2	B6165
15	35	60.3	1701	2	T09127
16	34	58.6	281	2	G72680
17	34	58.6	317	2	T39736
18	34	58.6	327	2	B91181
19	34	58.6	327	2	B61404
20	34	58.6	327	2	A86028
21	34	58.6	386	2	T24687
22	34	58.6	386	2	T24811
23	34	58.6	500	2	T19525
24	34	58.6	647	2	B64170
25	34	58.6	2793	2	B80784
26	34	58.6	2806	2	B85644
27	33	56.9	47	2	S22828
28	33	56.9	57	2	T35655
29	33	56.9	58	2	S10755

30	33	56.9	62	2	S10754	protamine Strb - h
31	33	56.9	66	2	B97818	30S ribosomal prot
32	33	56.9	113	2	S02365	hypothetical prote
33	33	56.9	115	2	T27446	hypothetical prote
34	33	56.9	159	1	W2BEE3	U14 protein - sui
35	33	56.9	285	2	C87585	hypothetical prote
36	33	56.9	333	1	A95846	probable ECF-sigma
37	33	56.9	405	2	D69812	ferriochrome ABC tr
38	33	56.9	406	2	E70303	hypothetical prote
39	33	56.9	408	2	B84743	hypothetical prote
40	33	56.9	410	2	T33489	hypothetical prote
41	33	56.9	441	2	B90051	hypothetical prote
42	33	56.9	467	2	AC2390	transposase all467
43	33	56.9	501	2	T45856	hypothetical prote
44	33	56.9	520	2	S14500	E2 glycoprotein pr
45	33	56.9	520	2	S14599	E2 glycoprotein pr

ALIGNMENTS

RESULT 1

T36275
hypothetical protein SCE68.25c - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T36275

R:Murphy, D.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A:Reference number: Z21576

A:Accession: T36275

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-52 <MDR>

A:Cross-references: UNIPROT:Q9WXX07, UNIPARC:UPI00000DB13D, EMBL:AL079345, PIDN:CA845361.

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOE68.25c

Query Match 72.4%; Score 42; DB 2; Length 32;
Best Local Similarity 75.0%; Pred. No. 0.63;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RKMKSTRRQRR 12
Db 18 RKMKSTRRQRR 29

RESULT 2

A87213
hypothetical protein ML2428A [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C:Accession: A87213

R:Coile, S.T.; Bigdelier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho

R.; Davies, R.M.; Devlin, K.; Dutnoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,

eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A:Authors: Rutherford, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; PMID:11234002

A:Accession: A87213

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-33 <STO>

A:Cross-references: UNIPROT:Q9CB56, UNIPARC:UPI0000139A1B, GB:AL450380, NID:G13094003, P

C:Genetics:

A:Gene: ML2428A

Query Match 67.2%; Score 39; DB 2; Length 33;
Best Local Similarity 66.7%; Pred. No. 2.2;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RKMUKSTRRORR 12
||:|:|:|
Db 18 RKLRRTRVQR 29

RESULT 3

T24499
hypotheoretical protein T05A6.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T24499

R:Thomas, K.
submitted to the EMBL Data Library, August 1995

A:Accession: T24499

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-258 <WIL>

A:Cross-references: UNIPROT:Q22198; UNIPARC:UPI000007A24C; EMBL:Z50796; PIDN:CAA90670.1;

C:Genetics:

A:Gene: CESP:T05A6.2

A:Map position: 2

A:introns: 25/3; 134/2; 157/1; 211/2

Query Match 67.2%; Score 39; DB 2; Length 258;

Best Local Similarity 66.7%; Pred. No. 14;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 RKMUKSTRRORR 12
||:|:|:|
Db 240 KKMTRSTRSR 251

RESULT 4

AE1979
ABC transporter ATP-binding protein alr1384 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2004

C:Accession: AE1979

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Itiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE1979

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-361 <KUR>

A:Cross-references: UNIPROT:Q8YX34; UNIPARC:UPI00000CE09F; GB:BA000019; PIDN:BAW73341.1;

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr1384

Query Match 65.5%; Score 38; DB 2; Length 361;

Best Local Similarity 72.7%; Pred. No. 29;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RKMUKSTRRORR 11
||:|:|:|
Db 122 RMOKSDRRQR 132

RESULT 5

T10219
protein kinase homolog T30C3.60 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C:Accession: T10219

R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16897

A:Accession: T10219
A:Molecule type: DNA
A:Residues: 1-651 <BEV>
A:Cross-references: UNIPROT:Q9STJ8; UNIPARC:UPI00000489E8; EMBL:AL079350; GSPDB:GN00062;
A:Experimental source: cultivar Columbia; BAC clone T30C3
C:Genetics:
A:Gene: ATSP:T30C3.60
A:Map position: 4

Query Match 63.8%; Score 37; DB 2; Length 651;

Best Local Similarity 77.8%; Pred. No. 76;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 4 LKSTRRQR 12
||:|:|:|
Db 643 LKTRRQR 651

RESULT 6

S39424

protamine P1 - Australian echidna

C:Species: Tachygonus aculeatus (Australian echidna)

C:Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004

C:Accession: S39424

R:Retief, J.D.; Winkler, R.J.; Dixon, G.H.

Eur. J. Biochem. 218, 457-461, 1993

A:Title: Evolution of the monoterpenes. The sequences of the protamine P1 genes of platyru

A:Reference number: S39424; MUID:94094837; PMID:8269934

A:Accession: S39424

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-69 <RET>

A:Cross-references: UNIPROT:P35311; UNIPARC:UPI000016C71A; EMBL:Z26848; NID:9407183; PIDN:

C:Genetics:

A:introns: 53/1

C:Superfamily: sperm histone

C:Keywords: DNA binding

Query Match 62.1%; Score 36; DB 2; Length 69;

Best Local Similarity 50.0%; Pred. No. 15;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RKMUKSTRRORR 12
||:|:|:|
Db 54 RSMRSTRRRR 65

RESULT 7

T29995
hypotheoretical protein C43H6.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T29995

R:Le, T.T.
submitted to the EMBL Data Library, March 1996

A:Description: The sequence of C. elegans coeimid C43H6.

A:Reference number: Z20717

A:Accession: T29995

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-89 <LET>

A:Cross-references: UNIPROT:Q18590; UNIPARC:UPI000007B96; EMBL:U51999; PIDN:AAA96089.1;

A:Experimental source: strain Bristol N2; clone C43H6

C:Genetics:

A:Gene: CESP:C43H6.8

A:Map position: X

A:introns: 42/2

Query Match 62.1%; Score 36; DB 2; Length 89;

Best Local Similarity 58.3%; Pred. No. 18;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RKMUKSTRRORR 12

Db 16 RKSKAERRKR 27

RESULT 8

S15153 hypothetical protein 110 - phage Pfl

C:Species: phage Pfl

C>Date: 18-Feb-1994 #sequence_revision 19-Jan-1996 #text_change 09-Jul-2004

C:Accession: S15153; #20824

R:Hall, D.R.; Short, N.U.; Perham, R.N.; Petersen, G.B.

J. Mol. Biol. 218, 349-364, 1991

A:Title: DNA sequence of the filamentous bacteriophage Pfl.

A:Reference number: S15140; MUID:91186399; PMID:2010913

A:Accession: S15153

A:Molecule type: DNA

A:Residues: 1-110 <JMO>

A:Cross-references: UNIPROT:P25132; UNIPARC:UPI00001384CE; EMBL:X52107; NID:g14829; PIDN

A:Experimental source: ATCC 25102-B1

Query Match 60.3%; Score 35; DB 2; Length 110;

Best Local Similarity 66.7%; Pred. No. 33;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 1 RKMLSTRQR 12

68 RPALSTRQR 79

RESULT 9

A85904 probable transcription regulator LYSR-type yf1E [imported] - Escherichia coli (strain O1

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 05-Oct-2004

C:Accession: A85904

P:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

11ler, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85460; MUID:21074935; PMID:11206551

A:Accession: A85904

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <STO>

A:Cross-references: UNIPROT:Q8X5A8; UNIPARC:UPI000000DCE5; GB:AE005174; NID:g12516998; F

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: yf1E

C:Superfamily: HTH-type transcriptional regulator, LysR family, MetR type

Query Match 60.3%; Score 35; DB 2; Length 308;

Best Local Similarity 58.3%; Pred. No. 86;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 1 RKMLSTRQR 12

213 RQFESTLRQR 224

RESULT 10

H65035 hypothetical protein emb-ung intergenic region - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 05-Oct-2004

C:Accession: H65035

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perma, N.T.; Burland, V.; Riley, M.; Co

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:9742617; PMID:9278503

A:Accession: H65035

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Residues: 1-308 <BLAT>

A:Cross-references: UNIPARC:UPI0000168148; GB:AE000344; GB:U00096; NID:g1788927; PIDN:AA

A:Experimental source: strain K-12, substrain M61655

C:Genetics:

A:Gene: yf1E

C:Superfamily: HTH-type transcriptional regulator, LysR family, MetR type

Query Match 60.3%; Score 35; DB 2; Length 308;

Best Local Similarity 58.3%; Pred. No. 86;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 1 RKMLSTRQR 12

213 RQFESTLRQR 224

RESULT 11

C91059 probable transcription regulator LYSR-type [imported] - Escherichia coli (strain O157:H7

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 05-Oct-2004

C:Accession: C91059

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kunihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

NA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: C91059

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <HNY>

A:Cross-references: UNIPROT:Q8X5A8; UNIPARC:UPI000000DCE5; GB:BA000007; PIDN:BA036866.1;

A:Experimental source: strain O157:H7, substrain RMD 050952

C:Genetics:

A:Gene: EC83443

C:Superfamily: HTH-type transcriptional regulator, LysR family, MetR type

Query Match 60.3%; Score 35; DB 2; Length 308;

Best Local Similarity 58.3%; Pred. No. 86;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 1 RKMLSTRQR 12

213 RQFESTLRQR 224

RESULT 12

A40894 RNA-directed RNA polymerase (EC 2.7.7.48) - Yeast (Saccharomyces cerevisiae) RNA replicor

C:Species: Saccharomyces cerevisiae

C>Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 05-Oct-2004

C:Accession: A40894; A40895

R:Rodriguez-Cousino, N.; Esteban, L.M.; Esteban, R.

J. Biol. Chem. 266, 12772-12778, 1991

A:Title: Molecular cloning and characterization of W double-stranded RNA, a linear molecu

A:Reference number: A40894; MUID:91286317; PMID:2061340

A:Accession: A40894

A:Molecule type: genomic RNA

A:Residues: 1-829 <ROD>

A:Cross-references: UNIPROT:P25328; UNIPARC:UPI000134B87; GB:M63893; NID:g4090991; PIDN

R:Matsumoto, Y.; Winkler, R.B.

J. Biol. Chem. 266, 12779-12783, 1991

A:Title: Yeast 20 S RNA replicon. Replication intermediates and encoded putative RNA poly

A:Reference number: A40895; MUID:91286318; PMID:11648104

A:Accession: A40895

A:Molecule type: genomic RNA

A:Residues: 1-825 <V> <MAT>

A:Cross-references: UNIPARC:UPI000006BD1D; GB:M64034; NID:g335015; PID:g555371

C:Superfamily: RNA-directed RNA polymerase in w dsRNA

C:Keywords: nucleotidyltransferase

Query Match

Best Local Similarity 50.0%; Score 35; DB 2; Length 829;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 RKMKSTRRQRR 12
Db 320 RRLFKGLRRRR 331

RESULT 13

TJ1354
probable potassium channel α k chain 1 - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: TJ1354
R/Engelard, B.; Neu, A.; Ludwig, J.; Roeper, J.; Pongs, O.
submitted to the EMBL Data Library, July 1998
A/Description: Identification of three rat potassium channel genes homologous to D. melo
A/Reference number: 220983
A/Accession: TJ1354
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1017 <ENG>
A/Cross-references: UNIPROT:Q9R1T9; UNIPARC:UPI0000170A85; EMBL:AJ007628; NID:e1329997;
A/Experimental source: cortex
C/Genetics:
A/Gene: α k1
C/Keywords: potassium channel

Query Match 60.3%; Score 35; DB 2; Length 1017;
Best Local Similarity 77.8%; Pred. No. 2.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LKSTRRQRR 12
Db 166 LRSTRQRR 174

RESULT 14

E86165
P1SK9.2 protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: E86165
R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasey, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroes, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: E86165
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1020 <STO>
A/Cross-references: UNIPROT:Q9ZVT9; UNIPARC:UPI000017CB88; GB:AE005172; NID:g3850588; PI
C/Genetics:
A/Map position: 1

Query Match 60.3%; Score 35; DB 2; Length 1020;
Best Local Similarity 77.8%; Pred. No. 2.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LKSTRRQRR 12
Db 978 LKSTRQRR 986

RESULT 15

T09127
probable erythrocyte-binding protein MAEBL - Plasmodium yoelii

C/Species: Plasmodium yoelii
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: T09127

R/Kappe, S.H.T.; Noe, A.R.; Fraser, T.S.; Blair, P.L.; Adams, J.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 1230-1235, 1998
A/Title: A family of chimeric erythrocyte binding proteins of malaria parasites.
A/Reference number: 216577; MUID:98115903; PMID:9448314
A/Accession: T09127

A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA

A/Residues: 1-1701 <KAP>

A/Cross-references: UNIPROT:O61164; UNIPARC:UPI000007D433; EMBL:AF031886; NID:g2947227;

A/Experimental source: subspecies yoelii; strain YM

C/Genetics:

A/Gene: maeb1

A/Introns: 62/1; 1648/1; 1674/2; 1697/1

C/Keywords: alternative splicing; cell binding; erythrocyte invasion

Query Match 60.3%; Score 35; DB 2; Length 1701;
Best Local Similarity 58.3%; Pred. No. 4.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKMKSTRRQRR 12
Db 1354 KRLKRRQRR 1365

Search completed: February 4, 2006, 03:42:20
Job time : 24.6667 secs

Protein Sequence Searches - February 2005

All of the sequence databases on ABSs have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rwp) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW protein - protein search, using SW model

Run on: February 4, 2006, 03:26:53 ; Search time 135.556 Seconds
(without alignments)
62.457 Million cell updates/sec

Title: US-10-790-768a-1
Perfect score: 58
Sequence: 1 RKMKSTRQR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_prot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	74.1	32	Q4NDY7_9M1CC	Q4ndy7 arthrobacte
2	42	74.1	33	Q6AAZ0_PROAC	Q6AAZ0 proteobac
3	42	72.4	32	Q9WXX7_STRCO	Q9WXX7 streptomyce
4	42	72.4	33	Q4USX2_CORUK	Q4USX2 corynebacte
5	42	72.4	33	Q8FSG0_CORBF	Q8FSG0 corynebacte
6	42	72.4	33	Q8WTF5_CORGL	Q8WTF5 corynebacte
7	40	69.0	125	Q8FHB6_ECOL6	Q8FHB6 escherichia
8	40	69.0	589	Q9CTI1_CVACA	Q9CTI1 cyanidium c
9	40	69.0	646	Q6SUB3_MANSM	Q6SUB3 manheimia
10	39	67.2	33	Y2428_MYCLE	Y2428 mycobacteri
11	39	67.2	33	Y5004_MYCTU	P055C9 mycobacteri
12	39	67.2	69	RS21_TREDE	Q7JXR1 treponema d
13	39	67.2	75	Q6NJK5_CORDI	Q6NJK5 corynebacte
14	39	67.2	258	Q22198_CABEL	Q22198 caenorhabdi
15	39	67.2	259	Q9J6R5_CABEL	Q9J6R5 caenorhabdi
16	39	67.2	279	Q9AUV5_ORYSA	Q9AUV5 oryza sativ
17	39	67.2	383	Q4WQNZ_ASPFU	Q4WQNZ aspergillus
18	39	67.2	424	Q5VMZ9_ORYSA	Q5VMZ9 oryza sativ
19	39	67.2	674	Q7SBH2_NEUCR	Q7SBH2 neurospora
20	38	65.5	228	Q7ZWE7_BRARE	Q7ZWE7 brachydanto
21	38	65.5	361	Q8YX34_ANASP	Q8YX34 anabaena sp
22	38	65.5	655	Q512Y1_MAGGR	Q512Y1 magnaporthe
23	38	65.5	1830	Q4SJA7_TETNG	Q4SJA7 tetradon n
24	38	65.5	4417	Q582D2_TETNG	Q582D2 tetradon n
25	37	63.8	196	Q4RHP2_TETNG	Q4RHP2 tetradon n
26	37	63.8	282	Q8LHM0_ORYSA	Q8LHM0 oryza sativ
27	37	63.8	386	Q866L2_ORYSA	Q866L2 oryza sativ
28	37	63.8	402	KLF12_HUMAN	Q9Y4X4 homo sapien
29	37	63.8	402	KLF12_MOUSE	Q35738 mus musculu
30	37	63.8	402	Q8WMI3_HUMAN	Q8WMI3 homo sapien
31	37	63.8	402	Q5VZM7_HUMAN	Q5VZM7 homo sapien

32	37	63.8	402	2	Q6NWV9_MOUSE	Q6NWV9 mus musculu
33	37	63.8	502	2	Q51R74_MAGGR	Q51R74 magnaporthe
34	37	63.8	524	2	Q5ASA0_EMENT	Q5ASA0 aspergillus
35	37	63.8	651	2	Q9STUB_ARATH	Q9STUB arabidopsis
36	37	63.8	892	2	Q4QAU7_LETMA	Q4QAU7 leishmania
37	37	63.8	3303	2	Q81339_PLAF7	Q81339 plasmodium
38	36	62.1	68	1	H8P1_TACAC	Q622M3 caenorhabdi
39	36	62.1	89	2	Q622M3_CABER	Q622M3 caenorhabdi
40	36	62.1	89	2	Q18590_CABEL	Q18590 caenorhabdi
41	36	62.1	99	2	Q8H2R3_ORYSA	Q8H2R3 oryza sativ
42	36	62.1	115	2	Q610N3_BACAN	Q610N3 bacillus an
43	36	62.1	122	2	Q652U2_ORYSA	Q652U2 oryza sativ
44	36	62.1	127	2	P76887_ECOLI	P76887 escherichia
45	36	62.1	131	2	Q9CW37_MOUSE	Q9CW37 m mus muscu

ALIGNMENTS

```

RESULT 1
Q4NDY7_9M1CC PRELIMINARY; PRT; 32 AA.
AC Q4NDY7;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
GN Hypothetical protein.
OR Names=Arthrobact_1453;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteriae; Actinobacteriales;
OC Micrococcales; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Deter C., Glavina T.,
RA Hammon N., Ibrani S., Pluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Latimer P., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA00100010; EAL95472.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 32 AA; 4083 MW; 3DFEAD9F6371E689 CRC64;

Query Match 74.1%; Score 43; DB 2; Length 32;
Best Local Similarity 66.7%; Pred. No. 0.94;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RKMKSTRQR 12
|||:|
Db 18 RKLKTRHQR 29

RESULT 2
Q6AAZ0_PROAC PRELIMINARY; PRT; 33 AA.
AC Q6AAZ0;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
GN OrderedLocusNames=PPA0320;
OS Propionibacterium acnes.

```

```
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Propionibacteriaceae; Propionibacterium.
OX NCBI_TaxID=1747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KPA171202 / DSM 16379;
RX PubMed=15286373; DOI=10.1126/science.1100330;
RA Bruggemann H., Henne A., Hoster F., Liesegang H., Wielez A.,
RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.,
RT "The complete genome sequence of Propionibacterium acnes, a commensal
of human skin."
RL Science 305:671-673(2004).
KW EMBL; AE017283; AAT82076.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 33 AA; 4073 MW; 4256B042E26371E7 CRC64;

Query Match 74.1%; Score 43; DB 2; Length 33;
Best Local Similarity 75.0%; Pred. No. 0.97;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRORR 12
DB 18 RKLKRTTRVQRR 29

RESULT 3
Q9WKX07_STRCO PRELIMINARY; PRT; 32 AA.
ID Q9WKX07_STRCO
AC Q9WKX07;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SC033327.
GN OrderedLocusNames=SC033327; ORFNames=SC066.25C;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., O'Neill M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Lathe L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrett B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RT Nature 417:141-147(2002).
DR EMBL; AL939116; CAB45361.1; -; Genomic DNA.
DR PIR; T36275; T36275.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 32 AA; 4045 MW; 3BE18FA26371F5A9 CRC64;

Query Match 72.4%; Score 42; DB 2; Length 32;
Best Local Similarity 75.0%; Pred. No. 1.5;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRORR 12
DB 18 RKLKRTTRVQRR 29

RESULT 4
Q4JSX2_CORJK
ID Q4JSX2_CORJK PRELIMINARY; PRT; 33 AA.
AC Q4JSX2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
```

```
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=jk1904;
OS Corynebacterium jeikeium (strain K411).
OC Bacteria; Actinobacteridae; Actinobacteriaceae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=306537;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K411;
RX PubMed=15968079; DOI=10.1128/JB.187.13.4671-4682.2005;
RA Tauch A., Kaiser O., Hain T., Goessmann A., Weisshaar B.,
RA Albermeier A., Bekei T., Bischoff N., Brune I., Chakraborty T.,
RA Kalinowski J., Meyer F., Rupp O., Schneider S., Viehoveit P.,
RA Puhler A.;
RT "Complete Genome Sequence and Analysis of the Multiresistant
RT Nosocomial Pathogen Corynebacterium jeikeium K411, a Lipid-Requiring
RT Bacterium of the Human Skin Flora."
RL J. Bacteriol. 187:4671-4682(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K411;
RA Linke B., Tauch A.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR311997; CAI38085.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 33 AA; 4163 MW; 8156A587F163D4F5 CRC64;

Query Match 72.4%; Score 42; DB 2; Length 33;
Best Local Similarity 75.0%; Pred. No. 1.5;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRORR 12
DB 18 RKLKRTTRVQRR 29

RESULT 5
Q8FSG0_COREF PRELIMINARY; PRT; 33 AA.
ID Q8FSG0_COREF
AC Q8FSG0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CE0433;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteridae; Actinobacteriaceae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Ueda Y., Kikuchi E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens."
RL Genome Res. 13:1572-1579(2003).
DR EMBL; BA000035; BAC17243.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 33 AA; 4163 MW; 8156A587F163D4F5 CRC64;

Query Match 72.4%; Score 42; DB 2; Length 33;
Best Local Similarity 75.0%; Pred. No. 1.5;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRORR 12
DB 18 RKLKRTTRVQRR 29
```

RESULT 6

Q6NT95 CORCL PRELIMINARY; PRT; 33 AA.
ID Q6NT95 CORCL PRELIMINARY; PRT; 33 AA.
AC Q6NT95 CORCL PRELIMINARY; PRT; 33 AA.
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein Cg10413.
GN OrderedLocustNames=Cg10413, Cg0494;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteriales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
NCBI_Taxid=1718;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA MEDLINE=22830012; PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkhardt A., Duach N., Eggeling L., Eikmanns B.J., Gaisiglat L.,
RA Goessmann A., Hartmann M., Hutmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moschel B., Pfeifferle W., Puchler A.,
RA Rey D.A., Ruckert C., Rupp O., Sahm H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins";
RL J. Biotechnol. 104:5-25(2003).
DR EMBL; BA000036; BAB97806.1; -; Genomic DNA.
DR EMBL; BX927149; CAP19130.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 33 AA; 4163 MW; 8156587F163D4F5 CRC64;

Query Match 72.4%; Score 42; DB 2; Length 33;
Best Local Similarity 75.0%; Pred. No. 1.5;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RKMUKSTRRQRR 12
Db 18 RKMUKSTRRQRR 29

RESULT 7

Q8FHB6 ECOL6 PRELIMINARY; PRT; 125 AA.
ID Q8FHB6 ECOL6 PRELIMINARY; PRT; 125 AA.
AC Q8FHB6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative acid shock protein.
GN OrderedLocustNames=c1989;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_Taxid=217992;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O6:H1 / CPT073 / ATCC 700928 / UPEC;
RA MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackert J., Stroud D.,
RA Mayhew G.F., Robe D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Moley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

DR EMBL; AE016761; AAN80449.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 125 AA; 14911 MW; 46A66E9F7B435219 CRC64;

Query Match 69.0%; Score 40; DB 2; Length 125;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RKMUKSTRRQRR 12
Db 64 RKMUKSTRRQRR 75

RESULT 8

Q9ZTL1 CYACA PRELIMINARY; PRT; 589 AA.
ID Q9ZTL1 CYACA PRELIMINARY; PRT; 589 AA.
AC Q9ZTL1;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE RNA polymerase sigma factor.
GN Name=tpoD2;
OS Cyanidium caldarium.
OC Eukaryota; Rhodophyta; Bangiophyceae; Cyanidiales; Cyanidiaceae;
OC Cyanidium.
NCBI_Taxid=2771;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=II-D-2;
RA Troxler R.F., Tan S., Liu B.;
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF050634; AAD02575.1; -; Genomic_DNA.
DR HSP; Q9WY78; I1W7.
DR GO; GO:0016987; F:sigma factor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006352; P:transcription initiation; IEA.
DR InterPro; IPR009042; Sigma70_r1_2.
DR InterPro; IPR007627; Sigma70_r2.
DR InterPro; IPR007624; Sigma70_r3.
DR InterPro; IPR007630; Sigma70_r4.
DR InterPro; IPR009043; Sigma70.
DR Pfam; PF00140; Sigma70_r1_2; 1.
DR Pfam; PF04542; Sigma70_r2; 1.
DR Pfam; PF04539; Sigma70_r3; 1.
DR Pfam; PF04545; Sigma70_r4; 1.
DR PRINTS; PR00046; SIGMA70FCT.
DR PROSITE; PS00715; SIGMA70_1; UNKNOWN_1.
DR PROSITE; PS00716; SIGMA70_2; 1.
SQ SEQUENCE 589 AA; 69127 MW; 751E2631DCB9814F CRC64;

Query Match 69.0%; Score 40; DB 2; Length 589;
Best Local Similarity 58.3%; Pred. No. 79;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKMUKSTRRQRR 12
Db 92 RQVYKSSRKQRR 103

RESULT 9

Q6SUB3 MANSM PRELIMINARY; PRT; 646 AA.
ID Q6SUB3 MANSM PRELIMINARY; PRT; 646 AA.
AC Q6SUB3;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE tnp protein.
GN Name=tnp; OrderedLocustNames=MS0840;
OS Mannheimia succiniciproducens (Strain MBE155E).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
NCBI_Taxid=221988;

[1]
NM NUCLEOTIDE SEQUENCE.
RX PubMed=15378067; DOI=10.1038/abt1010;
RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
RT Kim C.H., Jeong H., Hur C.G., Kim J.U.;
RT "The genome sequence of the capnophilic rumen bacterium Mannheimia
succiniciproducens";
RL Nat. Biotechnol. 22:1275-1281(2004).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL, AB016827; AUAJ7447.1; -; Genomic_DNA.
DR GO, GO:0005524; F:ATP binding; IEA.
DR GO, GO:0016887; F:ATPase activity; IEA.
DR GO, GO:0000166; F:nucleotide binding; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transp_like.
DR Pfam, PF00005; ABC_tran; 2.
DR ProDom, PD000006; ABC_transporter; 1.
DR SMART, SM00382; AAA; 2.
DR PROSITE, PS00211; ABC_TRANSPORTER_1; UNKNOWN_2.
DR PROSITE, PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Complete proteome; Nucleotide-binding; Transport.
SQ SEQUENCE 646 AA; 73203 MW; 1C31B0A355CEBB6 CRC64;

Query Match 69.0%; Score 40; DB 2; Length 646;
Best Local Similarity 72.7%; Pred. No. 87;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 KMLKSTRRQR 12
Db 289 KMLRERRQR 299

RESULT 10

Y2428_MYCLE STANDARD; PRT; 33 AA.
AC P0A5D0; Q9CB56;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein ML2428.1
GN OrderedLocustNames=ML2428.1; ORFNames=ML2428A;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
NX NCBI_TaxID=1769;
RX [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TN;
RA MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagsels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus";
RL Nature 409:1007-1011(2001).

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC EMBL, AL589325; CAC31945.1; -; Genomic_DNA.
DR PIR, A87213; A87213.
DR LepToma; ML2428A; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 33 AA; 4145 MW; 8156A587F16291F5 CRC64;

Query Match 67.2%; Score 39; DB 1; Length 33;
Best Local Similarity 66.7%; Pred. No. 5.6;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KMLKSTRRQR 12
Db 18 KMLRTRVQR 29

RESULT 11

Y500A_MYCTU STANDARD; PRT; 33 AA.
AC P0A5G9; Q9CB56;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein RV0500.1/MT0521.1
GN OrderedLocustNames=RV0500.1, MT0521.1; ORFNames=RV0500B;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
NX NCBI_TaxID=1773;

[1]
NM NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37RV;
RA MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagsels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver K., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence";
RL Nature 393:537-544(1998).
RX [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A.L., Utermack T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bernal W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL J. Bacteriol. 184:5479-5490(2002).

[3]
RP IDENTIFICATION.
RC STRAIN=H37RV;
RX MEDLINE=22255591; PubMed=12368430;
RX Camus J.-C., Pryor M.J., Medigue C., Cole S.T.;
RT "Re-annotation of the genome sequence of Mycobacterium tuberculosis
RT H37RV";
RL Microbiology 148:2967-2973(2002).

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL, BX842573; CAB55291.1; -; Genomic_DNA.
DR EMBL, AB006953; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR TIGR, MT0521.1; -;
DR TubercuList; RV0500B; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 33 AA; 4145 MW; 8156A587F16291F5 CRC64;

Query Match 67.2%; Score 39; DB 1; Length 33;
 Best Local Similarity 66.7%; Pred. No. 5.6;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 RKMLKSTRROR 12
 Db 18 RKLLRRTRVQR 29

RESULT 12
 RS21_TREEDE STANDARD; PRT; 69 AA.
 AC 073JRL;
 DT 10-MAY-2005 (Rel. 47, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE 308 ribosomal protein S21.
 GN Name=TPU; OrderedLocuNames=TDE2473;
 OS Treponema denticola.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
 NCBI_TaxID=158;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=ATCC 35405 / DSM 14222;
 RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
 RA Seehardt R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
 Dodson R.J., Davidsen T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
 Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
 Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
 Gebregeorgis E., Geer K., Tsegaye G., Malek J.A., Ayodeji B.,
 Shatman S., McLeod M.P., Smajs D., Howell J.K., Pal S., Amin A.,
 Vaehtich P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
 Weinstein G.M., Norris S.U., Fraser C.M., Paulsen I.T.;
 RA "Completion of the genome of the oral pathogen Treponema denticola
 with other spirochete genomes."
 RT Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
 RL -1- SIMILARITY: Belongs to the ribosomal protein S21P family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AEO17254; AAS12990.1; -; Genomic_DNA.
 DR TIGR; TDE2473; -.
 DR HAMAP; MF 00358; -; 1.
 DR InterPro; IPR001911; Ribosomal_S21.
 DR Pfam; PF01165; Ribosomal_S21; 1.
 DR PRINTS; PR00976; RIBOSOMALS21.
 DR ProDom; PD005521; Ribosomal_S21; 1.
 DR Trifam; TIGR00030; S21P; 1.
 DR TIGRFAM; TIGR00030; S21P; 1.
 DR POSITE; PS01181; RIBOSOMAL_S21; 1.
 KM Complete proteome; Ribonucleoprotein; Ribosomal protein.
 SQ SEQUENCE 69 AA; 8502 MW; 0D0698BE6C4B055C CRC64;

Query Match 67.2%; Score 39; DB 1; Length 69;
 Best Local Similarity 63.6%; Pred. No. 12;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 RKMLKSTRROR 11
 Db 54 RKLMKTRRSR 64

RESULT 13
 O6NUKS_CORDI PRELIMINARY; PRT; 75 AA.
 AC O6NUKS;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE Hypothetical protein.
 GN OrderedLocuNames=DIP0396;
 OS Corynebacterium diphtheriae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacterium.
 NCBI_TaxID=1717;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Biotype gravis / NCTC 13129;
 RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
 RA Cerdano-Farraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,
 Pallen M.J., Bentley S.D., Beera G.S., Churcher C.M., James K.D.,
 De Zorja A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
 Hamlin N., Holroyd S., Jasegl K., Moule S., Quail M.A.,
 Rabinowitz E., Rutherford K.M., Thomson N.R., Unwin L.,
 Whitehead S., Barrall B.G., Parkhill J.;
 RA "The complete genome sequence and analysis of Corynebacterium
 RT diphtheriae NCTC13129."
 RL Nucleic Acids Res. 31:6516-6523(2003).
 DR EMBL; BX248355; CAE48900.1; -; Genomic_DNA.
 KM Complete proteome.
 SQ SEQUENCE 75 AA; 8862 MW; 2C7B5250B6915B2C CRC64;

Query Match 67.2%; Score 39; DB 2; Length 75;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 RKMLKSTRROR 12
 Db 60 RKLLRRTRVQR 71

RESULT 14
 Q22198_CABEL PRELIMINARY; PRT; 258 AA.
 ID Q22198 CABEL
 AC Q22198;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein Cki-2.
 GN Name=Cki-2; ORFNames=T05A6.2, T05A6.2a;
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; Z50796; CAA90670.1; -; Genomic_DNA.
 DR PIR; T24499; T24499.
 DR Ensemble; T05A6.2; Caenorhabditis elegans.
 DR WormBase; WBGene0000517; Cki-2.
 DR WormPep; T05A6.2a; CE18928.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004661; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
 DR GO; GO:0007050; P:cell cycle arrest; IEA.
 DR InterPro; IPR001175; CDI.
 DR Pfam; PF02234; CDI; 1.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 258 AA; 28523 MW; 06A1AF2400BDE4D6 CRC64;

Query Match 67.2%; Score 39; DB 2; Length 258;
 Best Local Similarity 66.7%; Pred. No. 50;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 RKMLKSTRROR 12
 Db 240 KKQWTTSTRRSR 251

```

RESULT 15
Q9U6R5 CAEEL PRELIMINARY; PRT; 259 AA.
ID Q9U6R5;
AC Q9U6R5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Cyclin-dependent kinase inhibitor (hypothetical protein ckl-2).
OS Name=ckl-2; ORFNames=T05A6.2, T05A6.2B;
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=N2 Bristol;
RX MEDLINE=20056449; PubMed=10587644; DOI=10.1038/70272;
RA Feng H., Zhong W., Punksody G., Gu S., Zhou L., Seabolt E.K.,
RA Kildrege E.T.;
RT "CtU-2 is required for the G1-to-S-phase transition and mitotic
RT chromosome condensation in Caenorhabditis elegans.";
RL Nat. Cell Biol. 1:486-492(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AF179359; AAF13869.1; -; mRNA.
DR Ensembl; T05A6.2; Caenorhabditis elegans.
DR WormBase; WBGene0000517; ckl-2.
DR WormPeP; T05A6.2b; CEJ1824.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007050; F:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
DR Complete proteome; Cyclin; Hypothetical protein; Kinase.
DR KW
SQ SEQUENCE 259 AA; 28651 MW; 904F9814A9897EA8 CRC64;

Query Match 67.2%; Score 39; DB 2; Length 259;
Best local Similarity 66.7%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKMLKSTRRORR 12
DB 241 KKMTTSTRSR 252

```

Search completed: February 4, 2006, 03:41:25
 Job time : 139.556 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 03:41:44 ; Search time 30.6667 Seconds
(without alignments)
32.351 Million cell updates/sec

Title: US-10-790-768a-1
Perfect score: 58
Sequence: 1 RKMLKSTRQR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 8265679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/6.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/H.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/RB.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	72.4	33	2	US-09-605-703B-2766
2	39	67.2	131	2	US-09-513-999C-8105
3	36	62.1	161	2	US-09-252-991A-27639
4	36	62.1	537	2	US-09-252-991A-25958
5	35	60.3	87	2	US-09-248-786A-25921
6	35	60.3	157	1	US-08-630-822A-60
7	35	60.3	157	1	US-09-005-069-60
8	35	60.3	157	1	US-09-171-156A-18
9	35	60.3	157	2	US-09-004-730A-18
10	35	60.3	157	2	US-08-981-799A-18
11	35	60.3	220	2	US-09-134-000C-3832
12	35	60.3	660	2	US-09-252-991A-22842
13	35	60.3	725	2	US-09-252-991A-23752
14	35	60.3	879	1	US-08-554-612C-1
15	34	58.6	67	2	US-09-270-767-60600
16	34	58.6	105	2	US-09-270-767-59291
17	34	58.6	125	2	US-09-270-767-45108
18	34	58.6	154	2	US-09-252-991A-20018
19	34	58.6	271	2	US-09-252-991A-23448
20	34	58.6	285	2	US-09-252-991A-18133
21	34	58.6	321	2	US-09-489-039A-10010
22	34	58.6	367	2	US-09-270-767-43888
23	34	58.6	514	2	US-09-800-729-124
24	34	58.6	575	2	US-10-104-047-2221
25	34	58.6	1038	2	US-09-081-385-151
26	34	58.6	1038	2	US-09-752-639-151
27	34	58.6	1038	2	US-09-712-813-151

28	34	58.6	1038	2	US-09-700-354A-151	Sequence 151, App
29	34	58.6	1745	2	US-09-800-729-89	Sequence 89, App1
30	33	56.9	154	2	US-09-248-796A-28156	Sequence 28156, A
31	33	56.9	204	2	US-09-252-991A-27153	Sequence 27153, A
32	33	56.9	247	2	US-09-270-767-32723	Sequence 32723, A
33	33	56.9	247	2	US-09-270-767-47940	Sequence 47940, A
34	33	56.9	267	2	US-09-252-991A-32147	Sequence 32147, A
35	33	56.9	287	2	US-09-270-767-32196	Sequence 32196, A
36	33	56.9	287	2	US-09-270-767-47413	Sequence 47413, A
37	33	56.9	326	2	US-09-252-991A-25470	Sequence 25470, A
38	33	56.9	612	2	US-09-252-991A-31248	Sequence 31248, A
39	33	56.9	612	2	US-09-902-540-10556	Sequence 10556, A
40	33	56.9	704	2	US-09-252-991A-17523	Sequence 17523, A
41	33	56.9	734	2	US-09-270-767-46596	Sequence 46596, A
42	33	56.9	1348	2	US-09-949-002-517	Sequence 517, App
43	33	56.9	2196	2	US-10-360-101-259	Sequence 259, App
44	33	56.9	2224	2	US-09-054-272-38	Sequence 38, App1
45	33	56.9	2224	2	US-09-949-002-292	Sequence 292, App

ALIGNMENTS

RESULT 1
US-09-605-703B-2766
Sequence 2766, Application US/09605703B
Patent No. 6962989
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Kroeger, Burkhard
APPLICANT: Schröder, Hartwig
APPLICANT: Zeidler, Oskar
APPLICANT: Heberhauser, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
FILE REFERENCE: BGI-129CP
CURRENT APPLICATION NUMBER: US/09/605, 703B
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: 60/142,764
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 60/152,318
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 2934
SEQ ID NO 2766
LENGTH: 33
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-605-703B-2766
Query Match 72.4%; Score 42; DB 2; Length 33;
Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Cy 1 RKMLKSTRQR 12
Db 18 RKMLKSTRQR 29
RESULT 2
US-09-513-999C-8105
Sequence 8105, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
Patent No. 6783961
FILE REFERENCE: 59 US2, PRC
CURRENT APPLICATION NUMBER: US/09/513, 999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26

```
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO: 8105
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 24
; OTHER INFORMATION: Xaa=Cys or Tyr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 107
; OTHER INFORMATION: Xaa= * or Cys or Trp
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 120
; OTHER INFORMATION: Xaa= * or Leu or Ser or Trp
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 121
; OTHER INFORMATION: Xaa=Met or Val
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 131
; OTHER INFORMATION: Xaa=Ala or Thr
US-09-513-999C-8105
```

```
Query Match 67.2%; Score 39; DB 2; Length 131;
Best Local Similarity 58.3%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy 1 RKMLKSTRRQR 12
Db 51 KQLKSRRRQRK 62
```

```
RESULT 3
US-09-252-991A-27639
; Sequence 27639, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27639
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27639
```

```
Query Match 62.1%; Score 36; DB 2; Length 161;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Oy 1 RKMLKSTRRQR 12
Db 2 RKAWKSSRRTR 13
```

```
RESULT 4
US-09-252-991A-25958
; Sequence 25958, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
```

```
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25958
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25958
```

```
Query Match 62.1%; Score 36; DB 2; Length 537;
Best Local Similarity 50.0%; Pred. No. 2,5e+02;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy 1 RKMLKSTRRQR 12
Db 402 RRLRTQRRRR 413
```

```
RESULT 5
US-09-248-796A-25921
; Sequence 25921, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 25921
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-25921
```

```
Query Match 60.3%; Score 35; DB 2; Length 87;
Best Local Similarity 58.3%; Pred. No. 71;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
Oy 1 RKMLKSTRRQR 12
Db 12 RKMLKKNLRRR 23
```

```
RESULT 6
US-08-630-822A-60
; Sequence 60, Application US/08630822A
; Patent No. 5840695
; GENERAL INFORMATION:
; APPLICANT: FRANK, GLENN R.
; APPLICANT: HUNTER, SHIRLEY WU
; APPLICANT: WALLENFELS, LYNDY
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
```

COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,822A
FILING DATE: 11-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CONNELL, GARY J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-630-822A-60

Query Match 60.3%; Score 35; DB 1; Length 157;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RKMLKSTRQ 10
||:|:|
Db 100 RKVLESYVRQ 109

RESULT 7
US-09-005-069-60
Sequence 60, Application US/09005069
Patent No. 5932470
GENERAL INFORMATION:
APPLICANT: FRANK, GLENN R.
APPLICANT: HUNTER, SHIRLEY WU
APPLICANT: WALENFELS, LYNDIA
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
City: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,069
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,822
FILING DATE: 11-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: CONNELL, GARY J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-005-069-60

Query Match 60.3%; Score 35; DB 1; Length 157;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RKMLKSTRQ 10
||:|:|
Db 100 RKVLESYVRQ 109

RESULT 8
US-09-171-156A-18
Sequence 18, Application US/09171156A
Patent No. 6368846
GENERAL INFORMATION:
APPLICANT: Hunter, Shirley Wu
Sim, Gek-Kee
Webster, Eric R.
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: SHERIDAN ROSS P.C.
STREET: 1560 BROADWAY, SUITE 1200
CITY: DENVER
STATE: CO
COUNTRY: U.S.A.
ZIP: 80202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,156A
FILING DATE: 04-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C4-FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-171-156A-18

Query Match 60.3%; Score 35; DB 2; Length 157;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RKMLKSTRQ 10
||:|:|
Db 100 RKVLESYVRQ 109

RESULT 9
US-09-004-730A-18
Sequence 18, Application US/09004730A
Patent No. 6485968
GENERAL INFORMATION:

APPLICANT: Weber, Eric
APPLICANT: Wu Hunter, Shitiley
APPLICANT: Sim, Gek-Kee
APPLICANT: Frank, Glenn
APPLICANT: Wallenfeld, Lynda
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND APPARATUS TO COLLECT SUCH
FILE REFERENCE: 2618-17-CS-PUS-1
CURRENT APPLICATION NUMBER: US/09/004,730A
CURRENT FILING DATE: 1998-01-08
PRIOR APPLICATION NUMBER: PCT/97US/18669
PRIOR FILING DATE: 1997-10-15
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn version 3.0
SEQ ID NO 18
LENGTH: 157
TYPE: PRT
ORGANISM: Ctenocephalides felis
US-09-004-730A-18

Query Match 60.3%; Score 35; DB 2; Length 157;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQ 10
DB 100 RKVLESVRQ 109

RESULT 10
US-08-981-799A-18
Sequence 18, Application US/08981799A
Patent No. 6576238
GENERAL INFORMATION:
APPLICANT: Weber, Eric
APPLICANT: Wu Hunter, Shitiley
APPLICANT: Sim, Gek-Kee
APPLICANT: Frank, Glenn
APPLICANT: Wallenfeld, Lynda
TITLE OF INVENTION: "NOVEL ECTOPARASITE SALIVA PROTEINS AND APPARATUS TO COLLECT SUCH
FILE REFERENCE: 2618-17-CS-PUS
CURRENT APPLICATION NUMBER: US/08/981,799A
CURRENT FILING DATE: 1998-08-27
PRIOR APPLICATION NUMBER: PCT/97/18669
PRIOR FILING DATE: 1997-10-15
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn version 3.0
SEQ ID NO 18
LENGTH: 157
TYPE: PRT
ORGANISM: Ctenocephalides felis
US-08-981-799A-18

Query Match 60.3%; Score 35; DB 2; Length 157;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQ 10
DB 100 RKVLESVRQ 109

RESULT 11
US-09-134-000C-3832
Sequence 3832, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3832
LENGTH: 220
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-3832

Query Match 60.3%; Score 35; DB 2; Length 220;
Best Local Similarity 87.5%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KSTRRRR 12
DB 1 KSTRRRR 8

RESULT 12
US-09-252-991A-22842
Sequence 22842, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22842
LENGTH: 660
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22842

Query Match 60.3%; Score 35; DB 2; Length 660;
Best Local Similarity 58.3%; Pred. No. 4.4e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RKMLKSTRRRR 12
DB 603 RRQLPQRRRR 614

RESULT 13
US-09-252-991A-23752
Sequence 23752, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23752
LENGTH: 725
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23752

Query Match 60.3%; Score 35; DB 2; Length 725;

Best Local Similarity 58.3%; Pred. No. 4.8e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQR 12
Db 647 RRQLPQARRRR 658

RESULT 14

US-08-554-612C-1
Sequence 1, Application US/08554612C
Patent No. 5747660
GENERAL INFORMATION:
APPLICANT: Orlicky, David
TITLE OF INVENTION: PROSTAGLANDIN P2 RECEPTOR REGULATORY
TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,612C
FILING DATE: No. 5747660ember 6, 1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: 06519/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 322-5070
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 879 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-554-612C-1

Query Match 60.3%; Score 35; DB 1; Length 879;

Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRQR 12
Db 860 KKEVETRRRR 871

RESULT 15

US-09-270-767-60600
Sequence 60600, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 60600
LENGTH: 67
TYPE: PRT
ORGANISM: Drosophila melanogaster

US-09-270-767-60600

Query Match 58.6%; Score 34; DB 2; Length 67;

Best Local Similarity 87.5%; Pred. No. 81;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKMLKSTR 8
Db 13 RKVLKSTR 20

Search completed: February 4, 2006, 03:43:33
Job time : 31.6667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 03:42:34 ; Search time 104 Seconds
(Without alignments)
48.211 Million cell updates/sec

Title: US-10-790-768A-1
Perfect score: 58
Sequence: 1 RKMLKSTRRQR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptcdatc/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptcdatc/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptcdatc/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptcdatc/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptcdatc/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptcdatc/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	12	4	US-10-790-768A-1
2	58	100.0	15	4	US-10-790-768A-2
3	58	100.0	15	4	US-10-790-768A-14
4	58	100.0	15	4	US-10-790-768A-16
5	58	100.0	21	4	US-10-790-768A-4
6	58	100.0	21	4	US-10-790-768A-19
7	58	100.0	21	4	US-10-790-768A-20
8	58	100.0	27	4	US-10-790-768A-5
9	48	82.8	19	4	US-10-790-768A-18
10	48	82.8	25	4	US-10-790-768A-21
11	48	82.8	25	4	US-10-790-768A-22
12	43	74.1	247	4	US-10-408-765A-2219
13	42	72.4	33	3	US-09-738-626-3964
14	40	69.0	121	4	US-10-437-963-113033
15	40	69.0	691	4	US-10-156-761-8521
16	39	67.2	33	4	US-10-080-170-279
17	39	67.2	33	4	US-10-080-170-405
18	39	67.2	33	4	US-10-080-170-279
19	39	67.2	33	4	US-10-080-170-405
20	39	67.2	33	4	US-10-468-356-279
21	39	67.2	33	4	US-10-468-356-405
22	39	67.2	129	4	US-10-767-701-62017
23	39	67.2	258	4	US-10-369-483-5460
24	38	65.5	50	5	US-10-741-600-1197
25	37	63.8	177	4	US-10-767-701-56381
26	37	63.8	193	4	US-10-767-701-58463
27	37	63.8	262	4	US-10-437-963-147451

28	37	63.8	282	4	US-10-437-963-122500	Sequence 122500,
29	37	63.8	402	5	US-10-723-670-4	Sequence 4, Appli
30	37	63.8	840	4	US-10-437-963-151528	Sequence 151528,
31	36	62.1	127	4	US-10-437-963-138917	Sequence 138917,
32	36	62.1	127	4	US-10-425-115-263163	Sequence 263163,
33	36	62.1	145	4	US-10-424-599-276649	Sequence 276649,
34	36	62.1	151	6	US-11-097-143-33525	Sequence 33525, A
35	36	62.1	156	4	US-10-767-701-36445	Sequence 36445, A
36	36	62.1	245	4	US-10-425-115-316007	Sequence 316007,
37	36	62.1	250	4	US-10-282-122A-46251	Sequence 46251, A
38	36	62.1	708	4	US-10-408-765A-1772	Sequence 1772, Ap
39	36	62.1	1301	4	US-10-425-115-239023	Sequence 239023,
40	35	60.3	59	4	US-10-425-115-349395	Sequence 349395,
41	35	60.3	62	4	US-10-425-115-190620	Sequence 190620,
42	35	60.3	66	4	US-10-437-963-108557	Sequence 108557,
43	35	60.3	79	4	US-10-424-599-220675	Sequence 220675,
44	35	60.3	86	4	US-10-424-599-145925	Sequence 145925,
45	35	60.3	153	4	US-10-437-963-141969	Sequence 141969,

ALIGNMENTS

RESULT 1
US-10-790-768A-1
: Sequence 1, Application US/10790768A
: Publication No. US20040209797A1
: GENERAL INFORMATION:
: APPLICANT: Karas, Michael
: TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
: FILE REFERENCE: 002877.00028
: CURRENT APPLICATION NUMBER: US/10/790, 768A
: CURRENT FILING DATE: 2004-03-03
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 1
: LENGTH: 12
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: protein transduction domain
US-10-790-768A-1

Query Match 100.0%; Score 58; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKMLKSTRRQR 12
Db 1 RKMLKSTRRQR 12

RESULT 2
US-10-790-768A-2
: Sequence 2, Application US/10790768A
: Publication No. US20040209797A1
: GENERAL INFORMATION:
: APPLICANT: Karas, Michael
: TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
: FILE REFERENCE: 002877.00028
: CURRENT APPLICATION NUMBER: US/10/790, 768A
: CURRENT FILING DATE: 2004-03-03
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 2
: LENGTH: 15
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: protein transduction domain
US-10-790-768A-2

Query Match 100.0%; Score 58; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKMLKSTRQR 12
Db 4 RKMLKSTRQR 15

RESULT 3
US-10-790-768A-14
; Sequence 14, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768A-14

Query Match 100.0%; Score 58; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKMLKSTRQR 12
Db 4 RKMLKSTRQR 15

RESULT 4
US-10-790-768A-16
; Sequence 16, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768A-16

Query Match 100.0%; Score 58; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKMLKSTRQR 12
Db 4 RKMLKSTRQR 15

RESULT 5
US-10-790-768A-4
; Sequence 4, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
US-10-790-768A-4

Query Match 100.0%; Score 58; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKMLKSTRQR 12
Db 4 RKMLKSTRQR 15

RESULT 6
US-10-790-768A-19
; Sequence 19, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768A-19

Query Match 100.0%; Score 58; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKMLKSTRQR 12
Db 4 RKMLKSTRQR 15

RESULT 7
US-10-790-768A-20
; Sequence 20, Application US/10790768A
; Publication No. US20040209797A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Karas, Michael
/ TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
/ FILE OF INVENTION: Acids
/ FILE REFERENCE: 002877.00028
/ CURRENT APPLICATION NUMBER: US/10/790.768A
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 20
/ LENGTH: 21
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: protein transduction domain
/ NAME/KEY: MISC FEATURE
/ LOCATION: (1)..(1)
/ OTHER INFORMATION: Bioclin
US-10-790-768a-20
```

```
Query Match          100.0%; Score 58; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RKMLKSTRRQR 12
Db      10 RKMLKSTRRQR 21
```

```
RESULT 8
US-10-790-768a-5
/ Sequence 5, Application US/10790768A
/ Publication No. US20040209797A1
/ GENERAL INFORMATION:
/ APPLICANT: Karas, Michael
/ TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
/ FILE OF INVENTION: Acids
/ FILE REFERENCE: 002877.00028
/ CURRENT APPLICATION NUMBER: US/10/790.768A
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5
/ LENGTH: 27
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: protein transduction domain
US-10-790-768a-5
```

```
Query Match          100.0%; Score 58; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RKMLKSTRRQR 12
Db      10 RKMLKSTRRQR 21
```

```
RESULT 9
US-10-790-768a-18
/ Sequence 18, Application US/10790768A
/ Publication No. US20040209797A1
/ GENERAL INFORMATION:
/ APPLICANT: Karas, Michael
/ TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
/ FILE OF INVENTION: Acids
/ FILE REFERENCE: 002877.00028
/ CURRENT APPLICATION NUMBER: US/10/790.768A
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: PatentIn version 3.1
```

```
/ SEQ ID NO 18
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: protein transduction domain
/ NAME/KEY: MISC FEATURE
/ LOCATION: (1)..(1)
/ OTHER INFORMATION: Bioclin
US-10-790-768a-18
```

```
Query Match          82.8%; Score 48; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 MLKSTRRQR 12
Db      10 MLKSTRRQR 19
```

```
RESULT 10
US-10-790-768a-21
/ Sequence 21, Application US/10790768A
/ Publication No. US20040209797A1
/ GENERAL INFORMATION:
/ APPLICANT: Karas, Michael
/ TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
/ FILE OF INVENTION: Acids
/ FILE REFERENCE: 002877.00028
/ CURRENT APPLICATION NUMBER: US/10/790.768A
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 21
/ LENGTH: 25
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: protein transduction domain
/ NAME/KEY: MISC FEATURE
/ LOCATION: (1)..(1)
/ OTHER INFORMATION: Bioclin
US-10-790-768a-21
```

```
Query Match          82.8%; Score 48; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 MLKSTRRQR 12
Db      10 MLKSTRRQR 19
```

```
RESULT 11
US-10-790-768a-22
/ Sequence 22, Application US/10790768A
/ Publication No. US20040209797A1
/ GENERAL INFORMATION:
/ APPLICANT: Karas, Michael
/ TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
/ FILE OF INVENTION: Acids
/ FILE REFERENCE: 002877.00028
/ CURRENT APPLICATION NUMBER: US/10/790.768A
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 22
/ LENGTH: 25
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
```

OTHER INFORMATION: protein transduction domain
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1)..(1)
OTHER INFORMATION: Biotin
US-10-790-768a-22

Query Match 82.8%; Score 48; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 0 29;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MKSTRRRR 12
Db 16 MKSTRRRR 25

RESULT 12
US-10-408-765A-2219
Sequence 2219, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fany, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
FILE REFERENCE: 660088.465
CURRENT APPLICATION NUMBER: US/10/408,765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2219
LENGTH: 247
TYPE: PRT
ORGANISM: Homo sapiens
US-10-408-765A-2219

Query Match 74.1%; Score 43; DB 4; Length 247;
Best Local Similarity 81.8%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RKMLKSTRRR 11
Db 130 RKSLKSTRRR 140

RESULT 13
US-09-738-626-3964
Sequence 3964, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENO, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-11-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 3964
LENGTH: 33
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-3964

Query Match 72.4%; Score 42; DB 3; Length 33;
Best Local Similarity 75.0%; Pred. No. 4.1;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRR 12
Db 18 RKLRRTRVQR 29

RESULT 14
US-10-437-963-113033
Sequence 113033, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 113033
LENGTH: 121
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_16860C.1.pep
US-10-437-963-113033

Query Match 69.0%; Score 40; DB 4; Length 121;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKMLKSTRRR 12
Db 107 RRLRRRRR 118

RESULT 15
US-10-156-761-8521
Sequence 8521, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30

Mon Feb 6 10:11:04 2006

us-10-790-768a-1.rapbm

Page 5

: PRIOR APPLICATION NUMBER:JP 2001-272697
 : PRIOR FILING DATE: 2001-08-02
 : NUMBER OF SEQ ID NOS: 15109
 : SEQ ID NO 8521
 : LENGTH: 691
 : TYPE: PRT
 : ORGANISM: Streptomyces avermitilis
 US-10-156-761-8521

Query Match	69.0%	Score 40	DB 4	Length 691
Best Local Similarity	66.7%	Pred. No. 1.9e+02		
Matches	8	Conservative	2	Mismatches 0
				Gaps 0
Qy	1	RKMLKSTRORR	12	
		: :		
Db	359	RKRLSRRRQRK	370	

Search completed: February 4, 2006, 03:47:33
Job time : 105 secs

THIS PAGE BLANK (USP 10)

CC represents PTD peptide used in an experiment to determine the
transduction efficiency of inverted isomers of PTDs.

XX Sequence 15 AA;

Query Match 100.0%; Score 75; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KGRKMLKSTRRQR 15
DB 1 KGRKMLKSTRRQR 15

RESULT 2

ADSI17618

ID ADS17618 standard; peptide; 15 AA.

AC ADS17618;

DT 02-DEC-2004 (first entry)

DE Amino acid sequence of protein transduction domain (PTD) peptide #2.

XX protein transduction domain; PTD; PTD-cargo moiety complex;
KW cell immortalisation; cell viability; internalising peptide;
KW cell transport.

XX Synthetic.

XX WO2004078933-A2.

XX 16-SEP-2004.

XX 04-MAR-2004; 2004WO-US006445.

XX 04-MAR-2003; 2003US-0451243P.

XX 03-MAR-2004; 2004US-00790768.

XX (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.

XX Karas M;

XX WPI; 2004-653708/63.

XX New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
intracellular compartment of a cell.

XX Claim 3; SEQ ID NO 2; 60pp; English.

XX The present sequence represents a peptide which functions as a protein
transduction domain (PTD), and is capable of delivering small molecules,
proteins, and nucleic acids to an intracellular compartment of a cell. An
amino terminal lysine linker improves the efficiency of the PTD. The PTD
can be used in PTD-cargo moiety complexes that can reversibly immortalise
cells and increase cell viability in culture. The present PTD has a
lysine linker.

XX Sequence 15 AA;

Query Match 100.0%; Score 75; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KGRKMLKSTRRQR 15
DB 1 KGRKMLKSTRRQR 15

RESULT 3

ADSI17636

ID ADS17636 standard; peptide; 21 AA.

XX ADS17636;

DT 02-DEC-2004 (first entry)

DE Amino acid sequence of a PTD designated peptide 15.

XX protein transduction domain; PTD; PTD-cargo moiety complex;
KW cell immortalisation; cell viability; internalising peptide;
KW cell transport.

XX Synthetic.

XX Key Location/Qualifiers
FT Modified-site 1
FT /note= "biotin attached"

XX WO2004078933-A2.

XX 16-SEP-2004.

XX 04-MAR-2004; 2004WO-US006445.

XX 04-MAR-2003; 2003US-0451243P.

XX 03-MAR-2004; 2004US-00790768.

XX (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.

XX Karas M;

XX WPI; 2004-653708/63.

XX New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
intracellular compartment of a cell.

XX Example 10; SEQ ID NO 20; 60pp; English.

XX The specification describes peptides which function as a protein
transduction domain (PTD), and are capable of delivering small molecules,
proteins, and nucleic acids to an intracellular compartment of a cell. An
amino terminal lysine linker improves the efficiency of the PTD. The PTD
can be used in PTD-cargo moiety complexes that can reversibly immortalise
cells and increase cell viability in culture. ADS17634-ADS17638 represent
PTD peptides used in an experiment to determine if introduction of a
nuclear localisation signal affects translocation of a PTD-cargo moiety
complex.

XX Sequence 21 AA;

Query Match 100.0%; Score 75; DB 8; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KGRKMLKSTRRQR 15
DB 7 KGRKMLKSTRRQR 21

RESULT 4

ADSI17620

ID ADS17620 standard; peptide; 21 AA.

AC ADS17620;

DT 02-DEC-2004 (first entry)

DE Amino acid sequence of protein transduction domain (PTD) peptide #3.

XX protein transduction domain; PTD; PTD-cargo moiety complex;
KW cell immortalisation; cell viability; internalising peptide;
KW cell transport.

OS Synthetic.
XX WO2004078933-A2.
XX
XX 16-SEP-2004.
XX
XX 04-MAR-2004; 2004WO-US006445.
XX
XX 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.
XX
XX (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.
XX
XX Karas M;
XX
XX WPI; 2004-653708/63.
XX
XX
XX New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
XX Disclosure; SEQ ID NO 4; 60pp; English.
XX
XX The present sequence represents a peptide which functions as a protein
CC transduction domain (PTD), and is capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. The present PTD has a
CC lysine linker and a nuclear localisation signal.
XX
SQ Sequence 21 AA;
Query Match 100.0%; Score 75; DB 8; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KGGKRMKSTRQR 15
DB 1 KGGKRMKSTRQR 15
RESULT 5
ADSI7635
ID ADSI7635 standard; peptide; 21 AA.
XX
XX ADSI7635;
XX
XX 02-DEC-2004 (first entry)
XX
XX Amino acid sequence of a PTD designated peptide 14.
XX
XX protein transduction domain; PTD; PTD-cargo moiety complex;
KW cell immortalisation; cell viability; internalising peptide;
KW cell transport.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "biotin attached"
PT
XX WO2004078933-A2.
XX
XX 16-SEP-2004.
XX
XX 04-MAR-2004; 2004WO-US006445.
XX
XX 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.
XX
XX (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.
XX
XX

PI Karas M;
XX
XX WPI; 2004-653708/63.
XX
XX New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
XX Example 10; SEQ ID NO 19; 60pp; English.
XX
XX The specification describes peptides which function as a protein
CC transduction domain (PTD), and are capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. ADSI7634-ADSI7638 represent
CC PTD peptides used in an experiment to determine if introduction of a
CC nuclear localisation signal affects translocation of a PTD-cargo moiety
CC complex.
XX
SQ Sequence 21 AA;
Query Match 100.0%; Score 75; DB 8; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KGGKRMKSTRQR 15
DB 1 KGGKRMKSTRQR 15
RESULT 6
ADSI7632
ID ADSI7632 standard; peptide; 15 AA.
XX
XX ADSI7632;
XX
XX 02-DEC-2004 (first entry)
XX
XX Amino acid sequence of a PTD designated Biotin-Invrp.
XX
XX protein transduction domain; PTD; PTD-cargo moiety complex;
KW cell immortalisation; cell viability; internalising peptide;
KW cell transport.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "biotin attached"
PT
XX WO2004078933-A2.
XX
XX 16-SEP-2004.
XX
XX 04-MAR-2004; 2004WO-US006445.
XX
XX 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.
XX
XX (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.
XX
XX Karas M;
XX
XX WPI; 2004-653708/63.
XX
XX New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
XX Example 4; SEQ ID NO 16; 60pp; English.
XX
XX The specification describes peptides which function as a protein

CC transduction domain (PTD), and are capable of delivering small molecules, proteins, and nucleic acids to an intracellular compartment of a cell. An amino terminal lysine linker improves the efficiency of the PTD. The PTD can be used in PTD-cargo moiety complexes that can reversibly immortalise cells and increase cell viability in culture. The present sequence CC represents PTD peptide used in an experiment to determine the transduction efficiency of inverted isomers of PTDs.

XX
SQ Sequence 15 AA;

Query Match 93.3%; Score 70; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GGRKMLKSTRQR 15
DB 2 GGRKMLKSTRQR 15

RESULT 7
ADSI7621
ID ADSI7621 standard; peptide; 27 AA.

XX
AC ADSI7621;

DT 02-DEC-2004 (first entry)

DE Amino acid sequence of protein transduction domain (PTD) peptide #4.

XX protein transduction domain; PTD; PTD-cargo moiety complex;
XX cell immortalisation; cell viability; internalising peptide;
KM cell transport.

XX
OS Synthetic.

XX
PN WO2004078933-A2.

XX
PD 16-SEP-2004.

XX
PF 04-MAR-2004; 2004WO-US006445.

XX
PR 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.

XX
PA (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.

XX
PI Karas M;

XX
PI WPI; 2004-653708/63.

XX
DR WPI; 2004-653708/63.

XX
PT New isolated and purified polypeptide with a protein transduction domain, useful for delivering small molecules, proteins and nucleic acids to an intracellular compartment of a cell.

XX
PS Disclosure; SEQ ID NO 5; 60pp; English.

XX
CC The present sequence represents a peptide which functions as a protein transduction domain (PTD), and is capable of delivering small molecules, proteins, and nucleic acids to an intracellular compartment of a cell. An amino terminal lysine linker improves the efficiency of the PTD. The PTD can be used in PTD-cargo moiety complexes that can reversibly immortalise cells and increase cell viability in culture. The present PTD has a lysine linker and 2 nuclear localisation signals.

XX
SQ Sequence 27 AA;

Query Match 82.7%; Score 62; DB 8; Length 27;
Best Local Similarity 71.4%; Pred. No. 0.0065;
Matches 15; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

OY 1 KGG-----KXMKSTRQR 15
DB 1 KGGKKRKVRKMLKSTRQR 21

RESULT 8
ADSI7634
ID ADSI7634 standard; peptide; 19 AA.

XX
AC ADSI7634;

DT 02-DEC-2004 (first entry)

DE Amino acid sequence of a PTD designated peptide 13.

XX
XX protein transduction domain; PTD; PTD-cargo moiety complex;
XX cell immortalisation; cell viability; internalising peptide;
KM cell transport.

XX
OS Synthetic.

XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "biotin attached"

XX
PN WO2004078933-A2.

XX
PD 16-SEP-2004.

XX
PF 04-MAR-2004; 2004WO-US006445.

XX
PR 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.

XX
PA (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.

XX
PI Karas M;

XX
PI WPI; 2004-653708/63.

XX
PT New isolated and purified polypeptide with a protein transduction domain, useful for delivering small molecules, proteins and nucleic acids to an intracellular compartment of a cell.

XX
PS Example 10; SEQ ID NO 18; 60pp; English.

XX
CC The specification describes peptides which function as a protein transduction domain (PTD), and are capable of delivering small molecules, proteins, and nucleic acids to an intracellular compartment of a cell. An amino terminal lysine linker improves the efficiency of the PTD. The PTD can be used in PTD-cargo moiety complexes that can reversibly immortalise cells and increase cell viability in culture. ADSI7634-ADSI7638 represent PTD peptides used in an experiment to determine if introduction of a nuclear localisation signal affects translocation of a PTD-cargo moiety complex.

XX
SQ Sequence 19 AA;

Query Match 80.0%; Score 60; DB 8; Length 19;
Best Local Similarity 73.7%; Pred. No. 0.0098;
Matches 14; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

OY 1 KGGK---MKSTRQR 15
DB 1 KGGKKRKVRKMLKSTRQR 19

RESULT 9
ADSI7638
ID ADSI7638 standard; peptide; 25 AA.

XX
AC ADSI7638;

DT 02-DEC-2004 (first entry)

DE Amino acid sequence of a PTD designated peptide 17.

XX proteol transduction domain; PTD; PTD-cargo moiety complex;
KM cell immortalisation; cell viability; internalising peptide;
KM cell transport.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "biotin attached"
XX
PN WO2004078933-A2.
XX
PD 16-SEP-2004.
XX
PF 04-MAR-2004; 2004WO-US006445.
XX
PR 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.
XX
PA (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.
XX
PI Karas M;
XX
DR WPI; 2004-653708/63.
XX
PT New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
PS Example 10; SEQ ID NO 22; 60pp; English.
XX
CC The specification describes peptides which function as a protein
CC transduction domain (PTD), and are capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. ADS17634-ADS17638 represent
CC PTD peptides used in an experiment to determine if introduction of a
CC nuclear localisation signal affects translocation of a PTD-cargo moiety
CC complex.
XX
SQ Sequence 25 AA;
XX
Query Match 80.0%; Score 60; DB 8; Length 25;
Best Local Similarity 73.7%; Pred. No. 0.013;
Matches 14; Conservative 1; Mismatches 0; Indels 4; Gaps 1;
QY 1 KGGRRK---MLKSTRRQR 15
Db 7 KGGKKRRKRLKSTRRQR 25
XX
RESULT 10
ADS17637
ID ADS17637 standard; peptide; 25 AA.
XX
AC ADS17637;
XX
DT 02-DEC-2004 (first entry)
XX
DE Amino acid sequence of a PTD designated peptide 16.
XX
KM protein transduction domain; PTD; PTD-cargo moiety complex;
KM cell immortalisation; cell viability; internalising peptide;
KM cell transport.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "biotin attached"
XX

PN WO2004078933-A2.
XX
PD 16-SEP-2004.
XX
PF 04-MAR-2004; 2004WO-US006445.
XX
PR 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.
XX
PA (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.
XX
PI Karas M;
XX
DR WPI; 2004-653708/63.
XX
PT New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
PS Example 10; SEQ ID NO 21; 60pp; English.
XX
CC The specification describes peptides which function as a protein
CC transduction domain (PTD), and are capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. ADS17634-ADS17638 represent
CC PTD peptides used in an experiment to determine if introduction of a
CC nuclear localisation signal affects translocation of a PTD-cargo moiety
CC complex.
XX
SQ Sequence 25 AA;
XX
Query Match 80.0%; Score 60; DB 8; Length 25;
Best Local Similarity 73.7%; Pred. No. 0.013;
Matches 14; Conservative 1; Mismatches 0; Indels 4; Gaps 1;
QY 1 KGGRRK---MLKSTRRQR 15
Db 1 KGGKKRRKRLKSTRRQR 19
XX
RESULT 11
ADS17617
ID ADS17617 standard; peptide; 12 AA.
XX
AC ADS17617;
XX
DT 02-DEC-2004 (first entry)
XX
DE Amino acid sequence of protein transduction domain (PTD) peptide #1.
XX
KM protein transduction domain; PTD; PTD-cargo moiety complex;
KM cell immortalisation; cell viability; internalising peptide;
KM cell transport.
XX
OS Synthetic.
XX
PN WO2004078933-A2.
XX
PD 16-SEP-2004.
XX
PF 04-MAR-2004; 2004WO-US006445.
XX
PR 04-MAR-2003; 2003US-0451243P.
PR 03-MAR-2004; 2004US-00790768.
XX
PA (BIOW-) BIOWHITTAKER TECHNOLOGIES INC.
XX
PI Karas M;
XX
DR WPI; 2004-653708/63.
XX

PT New isolated and purified polypeptide with a protein transduction domain,
PT useful for delivering small molecules, proteins and nucleic acids to an
PT intracellular compartment of a cell.
XX
XX
PS Claim 1; SEQ ID NO 1; 60bp; English.
XX
XX The present sequence represents a peptide which functions as a protein
CC transduction domain (PTD), and is capable of delivering small molecules,
CC proteins, and nucleic acids to an intracellular compartment of a cell. An
CC amino terminal lysine linker improves the efficiency of the PTD. The PTD
CC can be used in PTD-cargo moiety complexes that can reversibly immortalise
CC cells and increase cell viability in culture. The present sequence is a
CC reverse isomer of ADS17641, a previously identified internalising
CC peptide.
XX
XX
SQ Sequence 12 AA;
Query Match 77.3%; Score 58; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 RKMLKSTRRQR 15
| | | | | | | | | | | | | |
Db 1 RKMLKSTRRQR 12
RESULT 12
AAU50525
ID AAU50525 standard; protein; 63 AA.
XX
XX AAU50525;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX Propionibacterium acnes immunogenic protein #11421.
DE
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX
OS Propionibacterium acnes.
XX
XX WO200181581-A2.
PN
XX
XX 01-NOV-2001.
PD
XX
XX 20-APR-2001; 2001WO-US012865.
PE
XX
XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
PI
XX
XX WPI; 2001-616774/71.
DR
XX
XX N-PSDB; AAS59549.
DR
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
PT
XX
XX
XX Example 1; SEQ ID NO 11720; 1069bp; English.
XX
XX Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 63 AA;
Query Match 58.7%; Score 44; DB 4; Length 63;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 KGRKMLKSTRRQR 15
| | | | | | | | | | | | | |
Db 45 KGRKMLKSTRRQR 59
RESULT 13
ABM47044
ID ABM47044 standard; protein; 63 AA.
XX
XX ABM47044;
AC
XX
XX 20-OCT-2003 (first entry)
DT
XX
XX Propionibacterium acnes predicted ORF-encoded polypeptide #11720.
DE
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
XX
OS Propionibacterium acnes.
XX
XX WO2003033515-A1.
PN
XX
XX 24-APR-2003.
PD
XX
XX 11-OCT-2002; 2002WO-US032727.
PF
XX
XX 15-OCT-2001; 2001US-00978825.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Valiieve-Douglas J;
PI
XX
XX WPI; 2003-381789/36.
DR
XX
XX N-PSDB; ACF64478.
DR
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
PT
XX
XX
XX Example 1; SEQ ID NO 11720; 1481bp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM3624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared

CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridization. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 63 AA;

SO Query Match 58.7%; Score 44; DB 6; Length 63;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KCGRKM LKSTRRQR 15
Db 45 KGRKMLKSTRRQR 59

RESULT 14

AAG90210 ID AAG90210 standard; protein, 33 AA.

XX AAG90210;

DT 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 3964.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis.

XX Corynebacterium glutamicum.

XX EPI108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

XX N-PSDB; AAH65429.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying

XX mutation point of a gene, measuring expression of a gene, analyzing

XX expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 3964; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These

XX are useful for identifying the mutation point of a gene derived from a

XX mutant of coryneform bacterium, measuring expression amount and analyzing

CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office

XX Sequence 33 AA;

SO Query Match 57.3%; Score 43; DB 4; Length 33;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KCGRKM LKSTRRQR 15
Db 15 KGRKMLRSTRRQR 29

RESULT 15

ADJ70413 ID ADJ70413 standard; protein, 247 AA.

XX ADJ70413;

XX 06-MAY-2004 (first entry)

XX Human heat mitochondrial protein as a therapeutic target SeqID2219.

XX mitochondrial; human; screening assay; diabetes mellitus;

XX Huntington's disease; osteoarthritis;

XX Leber's hereditary optic neuropathy; LHON;

XX mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

XX myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

XX neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;

XX osteopathic; ophthalmological; cytostatic.

XX Homo sapiens.

XX WO2003087768-A2.

XX 23-OCT-2003.

XX 04-APR-2003; 2003MO-US010870.

XX 12-APR-2002; 2002US-0372843P.

XX 17-JUN-2002; 2002US-0389987P.

XX 20-SEP-2002; 2002US-0412418P.

XX (MITO-) MITOKOR.

XX (BUCK-) BUCK INST AGE RES.

XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;

XX Warnock DE;

XX WPI; 2003-845369/78.

XX Identifying a mitochondrial target for drug screening assays and for

XX treating diseases associated with altered mitochondrial function,

XX PT comprises detecting a modified polypeptide in a sample and correlating

XX with the disease.

XX Claim 1; SEQ ID NO 2219; 180bp; English.

XX This invention relates to novel mitochondrial targets that can be used

XX for therapeutic intervention in treating a disease associated with

XX altered mitochondrial function. Specifically, it refers to a method for

XX identifying proteins of the human heart mitochondrial proteome that are

XX useful for drug screening assays, as well as therapeutic targets. The

XX present invention describes a method for identifying such proteins that

XX can be used in the treatment of various diseases associated with altered

XX mitochondrial function including diabetes mellitus, Huntington's disease,

CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, neurotropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cytoskeletal activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.

XX
 SQ Sequence 247 AA;

Query Match 57.3%; Score 43; DB 7; Length 247;
 Best Local Similarity 81.8%; Pred. No. 71;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 RKMVKSTRROR 14
 |||||:
 Db 130 RKSLSKSTRRRR 140

Search completed: February 4, 2006, 03:36:14
 Job time : 313.056 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 03:36:41 ; Search time 28.333 Seconds
(without alignments)
50.938 Million cell updates/sec

Title: US-10-790-768a-2
Perfect score: 75
Sequence: 1 KGGKMKLSTRQRR 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	58.7	77	2 B40973	spermatid-specific
2	44	58.7	78	2 A40973	spermatid-specific
3	44	58.7	79	2 S56116	spermatid-specific
4	43	57.3	32	2 T36275	hypothetical prote
5	42	56.0	69	2 S39424	protamine p1 - Aus
6	41	54.7	1017	2 T31354	probable potassium
7	40	53.3	33	2 A87213	hypothetical prote
8	40	53.3	357	2 S43220	hypothetical prote
9	40	53.3	361	2 AE1979	ABC transporter AT
10	40	53.3	441	2 D95124	glycosyl transfera
11	40	53.3	441	2 F97994	conserved hypochet
12	39	52.0	44	2 AG1431	ribosomal protein
13	39	52.0	44	2 AE1805	ribosomal protein
14	39	52.0	118	2 S56117	spermatid-specific
15	39	52.0	258	2 T24499	hypothetical prote
16	39	52.0	419	2 S68803	probable transcrip
17	39	52.0	836	2 T21631	hypothetical prote
18	38	50.7	261	2 AG1596	ribonuclease H rnh
19	38	50.7	305	1 NKVLHM	core antigen - her
20	38	50.7	408	2 E84743	hypothetical prote
21	38	50.7	426	2 B65119	protein T2B6.2 (im
22	38	50.7	489	2 S51428	hypothetical prote
23	38	50.7	507	2 AC3036	glycerol-3-phospha
24	38	50.7	507	2 H98249	glpD gene homolog
25	38	50.7	556	2 C96037	probable ABC trans
26	38	50.7	934	2 H90195	leucyl-tRNA synthet
27	38	50.7	1020	2 E86165	ribosomal protein
28	37	49.3	44	2 C48396	protamine II-1 - p
29	37	49.3	45	2 B58208	

30	37	49.3	45	2 A90081	50S ribosomal prot
31	37	49.3	48	2 S73486	ribosomal protein
32	37	49.3	61	2 S39425	protamine p1 - duc
33	37	49.3	136	1 FOADH5	major core protein
34	37	49.3	143	1 HSUR6P	histone H2B-3, spe
35	37	49.3	148	1 HSUR8P	histone H2B-3, spe
36	37	49.3	198	1 FOAD72	major core protein
37	37	49.3	299	2 F89937	hypothetical prote
38	37	49.3	333	2 D70855	probable ilvc prot
39	37	49.3	400	2 T41806	LEF-9 orf62 - Bomb
40	37	49.3	516	2 G72857	late expression fa
41	37	49.3	651	2 T10219	protein kinase hom
42	37	49.3	761	2 T24230	hypothetical prote
43	37	49.3	3104	2 S20473	fatty-acid synthas
44	36.5	48.7	350	2 G71445	hypothetical prote
45	36.5	48.7	776	2 S41628	genome polypeptide

ALIGNMENTS

RESULT 1
B40973
spermatid-specific protein T2 precursor - common cuttlefish
N:Alternate names: arginine-rich protamine; testis-specific protein T2
C:Species: Sepia officinalis (common cuttlefish)
C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 09-Jul-2004
C:Accession: B40973; S14086
R:Motifs-Tyrou, D.; Charlier-Harlin, M.C.; Martin-Ponthieu, A.; Bouillon, C.; Van Dorst
J. Biol. Chem. 266, 17388-17395, 1991
A:Title: Cuttlefish spermatid-specific protein T. Molecular characterization of two vari
A:Reference number: A40973; MUID:91373359; PMID:1894625
A:Accession: B40973
A:Molecule type: protein
A:Residues: 1-77 <MOU>
A:Cross-references: UNIPROT:P80002; UNIPARC:UPI00001323AB
R:Martin-Ponthieu, A.; Motiers-Tyrou, D.; Belatche, D.; Sautiere, P.; Schindler, P.; van
Bur. J. Biochem. 195, 611-619, 1991
A:Title: Cuttlefish sperm protamines. 1. Amino acid sequences of two distinct variants.
A:Reference number: S14085; MUID:9153296; PMID:1999185
A:Accession: S14086
A:Molecule type: protein
A:Residues: 22-77 <MAR>
A:Cross-references: UNIPARC:UPI00001771P9
C:Superfamily: sperm histone
C:Keywords: DNA binding; nucleus; phosphoprotein; spermatogenesis
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-77/Product: protamine variant Sp2 #status experimental <MAT>

Query Match 58.7%; Score 44; DB 2; Length 77;
Best Local Similarity 53.3%; Pred. No. 2.3;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Cy 1 KGGKMKLSTRQRR 15
Db 19 KGGRRRRRRRRR 33

RESULT 2
A40973
spermatid-specific protein T1 precursor - common cuttlefish
N:Alternate names: arginine-rich protamine; testis-specific protein T1
C:Species: Sepia officinalis (common cuttlefish)
C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 09-Jul-2004
C:Accession: A40973; S14085
R:Motiers-Tyrou, D.; Charlier-Harlin, M.C.; Martin-Ponthieu, A.; Bouillon, C.; Van Dorst
J. Biol. Chem. 266, 17388-17395, 1991
A:Title: Cuttlefish spermatid-specific protein T. Molecular characterization of two vari
A:Reference number: A40973; MUID:91373359; PMID:1894625
A:Accession: A40973
A:Molecule type: protein
A:Residues: 1-78 <MOU>
A:Cross-references: UNIPROT:P80001; UNIPARC:UPI00001323A4

R;Martin-Pontheu, A.; Moutere-Tyrou, D.; Belaitche, D.; Sautiere, P.; Schindler, P.; van
Eur. J. Biochem. 195, 611-619, 1991
A;Title: Cuttlefish sperm proteins. 1. Amino acid sequences of two distinct variants.
A;Reference number: S14085; MUID:91153298; PMID:1999185
A;Accession: S14085
A;Molecule type: protein
A;Residues: 22-78 <MAR>
A;Cross-references: UNIPARC:UPI00001771F8
C;Superfamily: sperm histone
C;Keywords: DNA binding; nucleus; phosphoprotein; spermatogenesis
F;1-21/Domains: signal sequence #status predicted <SIG>
F;22-78/Product: protamine variant Spt #status experimental <MAT>

Query Match 58.7%; Score 44; DB 2; Length 78;
Best Local Similarity 53.3%; Pred. No. 2.3;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 KGGKMLKSTRRQR 15
| | | | | : : : : : | | | | |
Db 19 KGGRRRRRRRRRR 33

RESULT 3

S56116
Spermatid-specific protein T1 - longfin squid
C;Species: Loligo pealeii (longfin squid)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S56116
R;Moutere-Tyrou, D.; Martin-Pontheu, A.; Ledoux-Andula, N.; Kouach, M.; Jaquinod, M.; S
Biochem. J. 309, 529-534, 1995
A;Title: Squid spermatogenesis: molecular characterization of testis-specific pro-protam
A;Reference number: S56116; MUID:95351983; PMID:7626016
A;Accession: S56116
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-79 <MOU>
A;Cross-references: UNIPROT:Q7MAG6; UNIPARC:UPI00001771FC
C;Superfamily: sperm histone

Query Match 58.7%; Score 44; DB 2; Length 79;
Best Local Similarity 53.3%; Pred. No. 2.3;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 KGGKMLKSTRRQR 15
| | | | | : : : : : | | | | |
Db 19 KGGRRRRRRRRRR 33

RESULT 4

T36275
Hypothetical protein SCE68.25c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36275
R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, June 1999
A;Reference number: Z21576
A;Accession: T36275
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-33 <MUR>
A;Cross-references: UNIPROT:Q9W07; UNIPARC:UPI00000DB13D; EMBL:AL079345; PTDN:CAB45361.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOE68.SCE68.25c

Query Match 57.3%; Score 43; DB 2; Length 32;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 KGGKMLKSTRRQR 15
| | | | | : : : : : | | | | |
Db 15 KGGKMLKSTRRQR 29

RESULT 5

S39424
protamine P1 - Australian echidna
C;Species: Tachylosus aculeatus (Australian echidna)
C;Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C;Accession: S39424
R;Releif, J.D.; Winkfein, R.J.; Dixon, G.H.
Eur. J. Biochem. 218, 457-461, 1993

A;Title: Evolution of the monotremes. The sequences of the protamine P1 genes of platypu
A;Reference number: S39424; MUID:94094837; PMID:8269934
A;Accession: S39424

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-69 <RET>
A;Cross-references: UNIPROT:P35311; UNIPARC:UPI000016C71A; EMBL:Z26848; NID:9407183; PTD
C;Genetics:
A;Intons: 53/1
C;Superfamily: sperm histone
C;Keywords: DNA binding

Query Match 56.0%; Score 42; DB 2; Length 69;
Best Local Similarity 53.8%; Pred. No. 4.4;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 3 GRKMLKSTRRQR 15
| | | | | : : : : : | | | | |
Db 53 GRSMRRRRRRRR 65

RESULT 6

T31354
Probable potassium channel elk chain 1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31354
R;Engelard, B.; Neu, A.; Ludwig, J.; Roeper, J.; Pongs, O.
Submitted to the EMBL Data Library, July 1998
A;Description: Identification of three rat potassium channel genes homologous to D. melar

A;Reference number: Z20983
A;Accession: T31354
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1017 <ENG>
A;Cross-references: UNIPROT:Q9R1T9; UNIPARC:UPI0000170A85; EMBL:AJ007628; NID:el329997;
A;Experimental source: cortex
C;Genetics:
A;Gene: elk1
C;Keywords: potassium channel

Query Match 54.7%; Score 41; DB 2; Length 1017;
Best Local Similarity 53.3%; Pred. No. 74;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 KGGKMLKSTRRQR 15
| | | | | : : : : : | | | | |
Db 160 KASSRLRSTRRQR 174

RESULT 7

A87213
Hypothetical protein ML2428A [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87213
R;Cole, S.T.; Sigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthey, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutherford, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002

A/Accession: A67213
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-33 <STO>
A/Cross-references: UNIPROT:Q9CB56; UNIPARC:UPI0000139A1B; GB:AL450380; NID:gl3094003; F
C/Genetics:
A/Gene: ML2428A

Query Match 53.3%; Score 40; DB 2; Length 33;
Best Local Similarity 60.0%; Pred. No. 4.9;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KGGKMLKSTRRQR 15
Db 15 KGGKMLKSTRRQR 29

RESULT 8
S43220
hypothetical protein YER127w - yeast (Saccharomyces cerevisiae)

C/Species: Saccharomyces cerevisiae
C/Date: 28-May-1993 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004

C/Accession: S43220; S50630
R/Mulligan, J.T.; Dietrich, P.S.; Henessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor,
submitted to the EMBL Data Library, February 1993

A/Reference number: S30812

A/Accession: S43220

A/Molecule type: DNA

A/Residues: 1-357 <ML>
A/Cross-references: UNIPROT:P40079; UNIPARC:UPI0000531FB; GB:U18916; EMBL:L11119; NID:S
R/Dietrich, P.S.

submitted to the EMBL Data Library, December 1994

A/Description: The sequence of S. cerevisiae cosmid 9781, 8198, 9115, 9981, and lambda

A/Reference number: S50630

A/Accession: S50630

A/Molecule type: DNA

A/Residues: 1-357 <DIE>

A/Cross-references: UNIPARC:UPI0000531FB; EMBL:U18916; NID:gl384128; PIDN:AA03225.1; F
C/Genetics:
A/Gene: SGD:LCPS

A/Cross-references: SGD:S0000929

A/Map position: SR

Query Match 53.3%; Score 40; DB 2; Length 357;
Best Local Similarity 46.7%; Pred. No. 42;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KGGKMLKSTRRQR 15
Db 271 KGGKMLKSTRRQR 285

RESULT 9
AE1979
ABC transporter ATP-binding protein alr1384 [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120

C/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2004

C/Accession: AE1979
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A/Title: Complete Genome Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A/Reference number: AB1807; MUID:21595285; PMID:1175940

A/Accession: AE1979
A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-361 <KUR>
A/Cross-references: UNIPROT:Q8YX34; UNIPARC:UPI00000CE09F; GB:BA000019; PIDN:BA07341.1; F
C/Genetics:
A/Experimental source: strain PCC 7120

A/Gene: alr1384

Query Match 53.3%; Score 40; DB 2; Length 361;
Best Local Similarity 69.2%; Pred. No. 43;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 2 KGGKMLKSTRRQR 14
Db 120 KGGKMLKSTRRQR 132

RESULT 10

D95124
glycosyl transferase, group 1 SP1076 [imported] - Streptococcus pneumoniae (strain TIGR4

C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C/Accession: D95124

R/Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
nson, T.; Hickey, E.R.; Holt, I.E.

Science 293, 498-506, 2001

A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A/Reference number: A95000; MUID:21357209; PMID:11463916

A/Accession: D95124

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-441 <KUR>

A/Cross-references: UNIPROT:Q970X1; UNIPARC:UPI0000516CB; GB:AE005672; PIDN:AAK75189.1;
A/Experimental source: strain TIGR4

C/Genetics:
A/Gene: SP1076

Query Match 53.3%; Score 40; DB 2; Length 441;
Best Local Similarity 53.8%; Pred. No. 51;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KGGKMLKSTRRQ 13
Db 416 KGGKMLKSTRRQ 428

RESULT 11

P97994
conserved Hypothetical protein spr0982 [imported] - Streptococcus pneumoniae (strain R6)

C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C/Accession: P97994

R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
J. Bacteriol. 183, 5709-5717, 2001

A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A/Reference number: A97872; MUID:21429245; PMID:11544234

A/Accession: P97994

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-441 <KUR>
A/Cross-references: UNIPROT:Q8CWR6; UNIPARC:UPI00000E493B; GB:AE007317; PIDN:AAK99786.1;
C/Genetics:
A/Gene: spr0982

Query Match 53.3%; Score 40; DB 2; Length 441;
Best Local Similarity 53.8%; Pred. No. 51;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KGGKMLKSTRRQ 13
Db 416 KGGKMLKSTRRQ 428

RESULT 12

AG1431
ribosomal protein L34 [imported] - Listeria monocytogenes (strain EGD-e)

C/Species: *Listeria monocytogenes*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AG1431
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Feihl, H.; D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A./Title: Comparative genomics of *Listeria species*.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AG1431
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-44 <GLA>
A/Cross-references: UNIPROT:Q926Q3; UNIPARC:UPI00000555E0; GB:NC_003210; PIDN:CAD01069.1
A/Experimental source: strain BGD-e
C/Genetics:
A/Gene: rpmH
C/Superfamily: *Escherichia coli* ribosomal protein L34

Query Match 52.0%; Score 39; DB 2; Length 44;
Best Local Similarity 53.3%; Pred. No. 9.3;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KGGKMLKSTRROR 15
DB 25 KNGRRVLASRRKGR 39

RESULT 13
AE1805
ribosomal protein L34 (imported) - *Listeria innocua* (strain Clijp11262)
C/Species: *Listeria innocua*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AE1805
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Feihl, H.; D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A./Title: Comparative genomics of *Listeria species*.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AE1805
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-44 <GLA>
A/Cross-references: UNIPROT:Q926Q3; UNIPARC:UPI00000555E0; GB:AL592022; PIDN:CA98213.1
A/Experimental source: strain Clijp11262
C/Genetics:
A/Gene: rpmH
C/Superfamily: *Escherichia coli* ribosomal protein L34

Query Match 52.0%; Score 39; DB 2; Length 44;
Best Local Similarity 53.3%; Pred. No. 9.3;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KGGKMLKSTRROR 15
DB 25 KNGRRVLASRRKGR 39

RESULT 14
S56117
spermatid-specific protein T2 precursor - longfin squid
N/Alternate names: sperm protamin SP
C/Species: *Loligo pealeii* (longfin squid)
C/Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C/Accession: S56117
R/Moutere-Tyrou, D.; Martin-Ponchieu, A.; Ledoux-Andula, N.; Kouach, M.; Jaquinod, M.; S Biochem. J. 309, 529-534, 1995
A/Title: Squid spermiogenesis: molecular characterization of testis-specific pro-protamin

A/Reference number: S56116; MUID:95351983; PMID:7626016
A/Accession: S56117
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-118 <MOU>
A/Cross-references: UNIPROT:Q7M4A3; UNIPARC:UPI000017BDD7

Query Match 52.0%; Score 39; DB 2; Length 118;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KGGKMLKSTRROR 14
DB 19 KGGRRRRRRRRRR 32

RESULT 15
T24499
hypothetical protein T05A6.2 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T24499
R/Thomas, K.
submitted to the EMBL Data Library, August 1995
A/Reference number: Z19899
A/Accession: T24499
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-258 <WTL>
A/Cross-references: UNIPROT:Q22198; UNIPARC:UPI000007A24C; EMBL:Z50796; PIDN:CAA90670.1
A/Experimental source: clone T05A6
C/Genetics:
A/Gene: CESP:T05A6.2
A/Map position: 2
A/Introns: 25/3; 134/2; 157/1; 211/2

Query Match 52.0%; Score 39; DB 2; Length 258;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 KXMLKSTRROR 15
DB 240 KXMTSTRSRNR 251

Search completed: February 4, 2006, 03:42:21
Job time : 29.333 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: February 4, 2006, 03:26:53 ; Search time 169.444 Seconds
(without alignments)
62.457 Million cell updates/sec

Title: US-10-790-768a-2

Sequence: 1 KGGKMKSTRRQR 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	47	62.7	QSR212_RAT	QSR212_rattus norv
2	45	60.0	Q7U8K9_SYNXP	Q7U8K9_synchococc
3	44	58.7	QANDY7_SMTCC	QANDY7_arthrobacte
4	44	58.7	Q6AAZ0_PPOAC	Q6AAZ0_propionibac
5	44	58.7	PRT1_SEPOF	P80002_sepia offic
6	44	58.7	PRT1_SEPOF	P80001_sepia offic
7	44	58.7	Q8C60 LOLOP	Q8C60_loligo opal
8	44	58.7	Q7M4G6_LOLPE	Q7M4G6_loligo opal
9	44	58.7	Q6ARV6_ORYSA	Q6ARV6_oryza sativ
10	44	58.7	Q8C0D6_MOUSE	Q8C0D6_mus muscicu
11	44	58.7	Q8C0F9_MOUSE	Q8C0F9_mus muscicu
12	44	58.7	Q8C0L5_MOUSE	Q8C0L5_mus muscicu
13	44	58.7	Q7P2L1_AMOGA	Q7P2L1_anopheles g
14	43	57.3	Q9WX07_SNRCC	Q9WX07_streptomyce
15	43	57.3	Q4USX2_CORUX	Q4USX2_corynebacte
16	43	57.3	Q8FSG0_COREL	Q8FSG0_corynebacte
17	43	57.3	Q8NT95_CORGL	Q8NT95_corynebacte
18	43	57.3	RLJ34_THEWA	P58288_thermocoga
19	43	57.3	Q4Q0T8_9SPHN	Q4Q0T8_erythroba
20	43	57.3	Q59JN6_CANAL	Q59JN6_candida alb
21	43	57.3	Q6ZIRS_ORYSA	Q6ZIRS_oryza sativ
22	43	57.3	Q8SY8_TETNG	Q8SY8_tetradodon n
23	43	57.3	Q8LMH0_ORYSA	Q8LMH0_oryza sativ
24	43	57.3	Q8S6L2_ORYSA	Q8S6L2_oryza sativ
25	43	57.3	Q4QAJ7_LEIMA	Q4QAJ7_leishmania
26	42	56.0	HSP1_TACAC	P33311_tachyglonu
27	42	56.0	Q8C37_MOUSE	Q8C37_m_mus muscu
28	42	56.0	Q5UJ3_BRARE	Q5UJ3_brachydiano
29	42	56.0	HEM4_PROMI	Q59683_protocus mir
30	42	56.0	Q7R3H5_GIALA	Q7R3H5_giardia lam
31	42	56.0	Q6RSF3_HPBUD	Q6RSF3_duck hepati

32	42	56.0	Q6RSF4_HPBUD	Q6RSF4_duck hepati
33	42	56.0	Q4KIC6_PSEF5	Q4KIC6_pseudomonas
34	42	56.0	Q13303_CRYCU	Q13303_cryptococcus
35	42	56.0	Q4RW3_TETNG	Q4RW3_tetradodon n
36	42	56.0	Q8BKX3_MOUSE	Q8BKX3_m_mus muscu
37	42	56.0	Q5ACY9_CANAL	Q5ACY9_candida alb
38	42	56.0	Q6FVZ6_CANGA	Q6FVZ6_candida gla
39	42	56.0	Q96JH4_HUMAN	Q96JH4_homo sapien
40	41	54.7	RK34_GRAVL	Q6B951_gracilaria
41	41	54.7	Q8FH6_ECOL6	Q8FH6_escherichia
42	41	54.7	Q4NW87_YDELTA	Q4NW87_aeromonas
43	41	54.7	Q69KAS_ORYSA	Q69KAS_oryza sativ
44	41	54.7	Q84ZK3_ORYSA	Q84ZK3_oryza sativ
45	41	54.7	Q89FAL_BRAVA	Q89FAL_bradynhixob

ALIGNMENTS

```
RESULT 1
QSR212_RAT PRELIMINARY; PRT; 406 AA.
AC QSR212;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Novel serine protease like protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yoshino M., Mizutani T., Yamada K., Yazawa T., Ogata H., Sekiguchi T.,
RA Kajitani T., Miyamoto K.;
RT "rat novel serine protease.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB180912; BAD74162.1; -; mRNA.
DR Ensembl; ENSRN000000025184; Rattus norvegicus.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1_S6.
DR Pfam; PF00089; Trypsin_1; UNKNOWN_1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydroxylase; Protease; Serine protease.
SQ SEQUENCE 406 AA; 45496 MW; 61F9C9A0E896EB2 CRC64;

Query Match 62.7%; Score 47; DB 2; Length 406;
Best Local Similarity 60.0%; Pred. No. 9.7;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KGGKMKSTRRQR 15
Db 191 KGGKMKSTRRQR 205

RESULT 2
Q7U8K9_SYNXP PRELIMINARY; PRT; 432 AA.
ID Q7U8K9;
AC Q7U8K9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Probable serine proteinase, periplasmic precursor (EC 3.4.21.-).
GN OrderedlocusNames=SYNM0608;
OS Synchococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=84588;
RN [1]
RP NUCLEOTIDE SEQUENCE.
MDLINB=22825697; PubMed=12917641; DOI=10.1038/nature01943;
```

RA Palenik B., Brahama B., Larimer F.W., Land M.L., Hauser L.,
 RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
 RA Paulsen I.T., Dufrene A., Paterinsky F., Webb E.A., Waterbury J.,
 RT "The genome of a motile marine Synechococcus".
 CC Nature 424:1037-1042(2003).
 RL -1- SIMILARITY: Contains 1 PDZ (DHR) domain.
 DR EMBL: BX569690; CAB07123.1; -; Genomic DNA.
 DR GO: GO:0008223; F:peptidase activity; IEA.
 DR GO: GO:0005515; F:protein binding; IEA.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR001940; Peptidase_S1C.
 DR InterPro: IPR001254; Peptidase_S1_S6.
 DR Pfam: PF00595; PDZ; 1.
 DR Pfam: PF00089; Trypsin; 1.
 DR PRINTS: PR00834; PROTEASES1C.
 DR SMART: SM00228; PDZ; 1.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS50106; PDZ; 1.
 DR Complete proteome; Hydrolase; Protease; Serine protease; Signal.
 KM SIGNAL
 FT SEQUENCE 432 AA; 46789 MW; 918E34B0CDF1942E CRC64;
 SQ
 Query Match 60.0%; Score 45; DB 2; Length 432;
 Best Local Similarity 64.3%; Pred. No. 24;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 KGGKMKLSTRRQR 14
 Db 13 KGGLELKHRRQR 26
 RESULT 3
 QANDY7_9M1CC
 ID QANDY7_9M1CC PRELIMINARY; PRT; 32 AA.
 AC QANDY7_9M1CC
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=ArthDRaft_1453;
 OS Arthrobacter sp. FB24.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcales; Micrococcaceae; Arthrobacter.
 OX NCBI_TaxID=290399;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=FB24;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Israni S., Pittluck S., Richardson P.,
 RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24."
 RN Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
 CC NUCLEOTIDE SEQUENCE.
 RC STRAIN=FB24;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Larimer F., Land M.,
 RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24."
 RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL: AAHQ100010; EAL95472.1; -; Genomic DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 32 AA; 4083 MW; 30FEAD9F6371E689 CRC64;
 Query Match 58.7%; Score 44; DB 2; Length 32;
 Best Local Similarity 60.0%; Pred. No. 2.2;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 KGGKMKLSTRRQR 15

Db 15 KGGKMKLSTRRQR 29
 RESULT 4
 Q6AAZ0_PROAC
 ID Q6AAZ0_PROAC PRELIMINARY; PRT; 33 AA.
 AC Q6AAZ0;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=PPA0320;
 OS Propionibacterium acnes.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Propionibacteriaceae; Propionibacterium.
 OX NCBI_TaxID=1747;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=KPA171202 / DSM 16379;
 RC PubMed=15286373; DOI=10.1126/science.1100330;
 RA Bruggemann H., Henne A., Hofer F., Liesegang H., Wierzer A.,
 RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.,
 RT "The complete genome sequence of Propionibacterium acnes, a commensal
 RT of human skin."
 RL Science 305:671-673(2004).
 DR EMBL: AE017283; AAT82076.1; -; Genomic DNA.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 33 AA; 4073 MW; 4256B042E26371E7 CRC64;
 Query Match 58.7%; Score 44; DB 2; Length 33;
 Best Local Similarity 66.7%; Pred. No. 2.3;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 KGGKMKLSTRRQR 15
 Db 15 KGGKMKLSTRRQR 29

RESULT 5
 PRT2_SEPOF
 ID PRT2_SEPOF STANDARD; PRT; 77 AA.
 AC P80002;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Spermaticid-specific protein T2 [Contains: Sperm protamine SP2].
 OS Sepia officinalis (Common cuttlefish).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Decapodiformes; Sepioidae; Sepiidae; Sepia.
 OX NCBI_TaxID=6610;
 RN PROTEIN SEQUENCE.
 RP MEDLINE=9137359; PubMed=1894625;
 RA Wouters-Tyrou D., Charlier-Harlin M.-C., Martin-Ponthieu A.,
 RA Boutillon C., Van Dorsselaer A., Sautiere P.,
 RT "Cuttlefish spermaticid-specific protein T2. Molecular characterization
 RT of two variants T1 and T2, putative precursors of sperm protamine
 RT variants Sp1 and Sp2."
 RL J. Biol. Chem. 266:17368-17395(1991).
 CC NUCLEOTIDE SEQUENCE OF 22-77.
 RP MEDLINE=91153298; PubMed=1999185;
 RA Martin-Ponthieu A., Wouters-Tyrou D., Belalche D., Sautiere P.,
 RA Schindler P., van Dorsselaer A.,
 RT "Cuttlefish sperm protamines. 1. Amino acid sequences of two distinct
 RT variants."
 RL Eur. J. Biochem. 195:611-619(1991).
 CC -1- FUNCTION: Cuttlefish spermatogenesis is characterized by a double
 CC nuclear protein transition: histones -> spermaticid-specific
 CC proteins (T1/T2) -> protamines (SP1/SP2). The protamines compact
 CC sperm DNA into a highly condensed, stable and inactive complex.
 CC -1- SUBCELLULAR LOCATION: Nuclear.

```
CC -1- TISSUE SPECIFICITY: Testis.
CC -1- DEVELOPMENTAL STAGE: Spermatogenesis.
CC -1- PPM: Phosphorylation occurs at different degrees. The
CC triphosphorylated form may be predominant in T2. SP2 appears to be
CC phosphorylated in elongated spermatids, but dephosphorylated in
CC mature sperm cells.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR PIR: B40973; B40973.
KW Chromosomal protein; Developmental protein; Differentiation;
KW Direct protein sequencing; DNA condensation; DNA-binding;
KW Multigene family; Nuclear protein; Nucleosome core; Phosphorylation;
KW Spermatogenesis; Testis.
FT CHAIN 1 77 Spermatid-specific protein T2.
FT REGION 22 77 Sperm protamine SP2.
FT COMPBIAS 1 21 Hydrophobic.
FT VARIANT 22 77 Arg-rich (highly basic).
FT VARIANT 1 1 Missing (in T2B).
SQ SEQUENCE 77 AA; 10485 MW; 0F2C1B215292ED07 CRC64;

Query Match 58.7%; Score 44; DB 1; Length 77;
Best Local Similarity 53.3%; Pred. No. 5.7;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KGGKRLKSTRRORR 15
Db 19 KGGRRRRRRRRRRR 33

RESULT 6
PRT1_SEPOP STANDARD; PRT; 78 AA.
ID PRT1_SEPOP STANDARD; PRT; 78 AA.
AC P80001;
DT 01-FEB-1991 (Rel. 20, Last sequence update)
DT 01-NOV-1991 (Rel. 47, Last annotation update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE Spermatid-specific protein T1 [Contains: Sperm protamine SP1].
OS Sepia officinalis (Common cuttlefish).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Decapodiformes; Sepioidae; Sepiidae; Sepia.
OC NCBI_Taxid=6610;
OX NCBI_Taxid=6610;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=91373359; PubMed=1894625;
RA Mouters-Tyrou D., Charlier-Harlin M.-C., Martin-Pontleu A.,
RA Boutillon C., van Dorselaer A., Sautiere P.,
RA "Cuttlefish spermatid-specific protein T1. Molecular characterization
RT of two variants T1 and T2, putative precursors of sperm protamine
RT variants Sp1 and Sp2."
RL J. Biol. Chem. 266:17388-17395 (1991).
RN [2]
RP PROTEIN SEQUENCE OF 22-78.
RX MEDLINE=91153296; PubMed=1999185;
RA Martin-Pontleu A., Mouters-Tyrou D., Belaiche D., Sautiere P.,
RA Schindler P., van Dorselaer A.;
RA "Cuttlefish sperm protamines. 1. Amino acid sequences of two distinct
RT variants."
RL Eur. J. Biochem. 195:611-619 (1991).
CC -1- FUNCTION: Cuttlefish spermatogenesis is characterized by a double
CC nuclear protein transition: histones -> spermatid-specific
CC protamine (T1/T2) -> protamines (SP1/SP2). The protamines compact
CC sperm DNA into a highly condensed, stable and inactive complex.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Testis.
CC -1- DEVELOPMENTAL STAGE: Spermatogenesis.
CC -1- PPM: Phosphorylation occurs at different degrees. The
CC triphosphorylated form may be predominant in T1. SP1 appears to be
CC phosphorylated in elongated spermatids, but dephosphorylated in
```

```
CC mature sperm cells.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR PIR: A40973; A40973.
KW Chromosomal protein; Developmental protein; Differentiation;
KW Direct protein sequencing; DNA condensation; DNA-binding;
KW Multigene family; Nuclear protein; Nucleosome core; Phosphorylation;
KW Spermatogenesis; Testis.
FT CHAIN 1 78 Spermatid-specific protein T1.
FT REGION 22 78 Sperm protamine SP1.
FT COMPBIAS 1 21 Hydrophobic.
FT COMPBIAS 22 78 Arg-rich (highly basic).
FT VARIANT 22 78 Arg-rich (highly basic).
FT VARIANT 1 1 Missing (in T1B).
SQ SEQUENCE 78 AA; 10632 MW; 09FE3EDBF0DCED33 CRC64;

Query Match 58.7%; Score 44; DB 1; Length 78;
Best Local Similarity 53.3%; Pred. No. 5.8;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KGGKRLKSTRRORR 15
Db 19 KGGRRRRRRRRRRR 33

RESULT 7
O86C60 LOLOP PRELIMINARY; PRT; 78 AA.
ID O86C60 LOLOP PRELIMINARY; PRT; 78 AA.
AC O86C60;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protamine.
OS Loligo opalescens (California market squid).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Decapodiformes; Loliginidae; Loligo.
OC NCBI_Taxid=31211;
OX NCBI_Taxid=31211;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15461424; DOI=10.1007/s00239-004-2589-8;
RA Lewis J.D., de Jong M.B., Bagha S.M., Tang A., Gilly W.F., Ausio J.;
RA "All roads lead to arginine: the squid protamine gene."
RL J. Mol. Evol. 58:673-680 (2004).
DR EMBL: AY269798; AAP32169.1; -; Genomic DNA.
SQ SEQUENCE 78 AA; 10651 MW; FD446437AF1B3278 CRC64;

Query Match 58.7%; Score 44; DB 2; Length 78;
Best Local Similarity 53.3%; Pred. No. 5.8;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KGGKRLKSTRRORR 15
Db 19 KGGRRRRRRRRRRR 33

RESULT 8
O7M4G6 LOLOP PRELIMINARY; PRT; 79 AA.
ID O7M4G6 LOLOP PRELIMINARY; PRT; 79 AA.
AC O7M4G6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Spermatid-specific protein T1.
OS Loligo pealeii (Longfin squid).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Decapodiformes; Loliginidae; Loligo.
OC NCBI_Taxid=6621;
OX NCBI_Taxid=6621;
RN [1]
RP PROTEIN SEQUENCE.
```

RA Moutere-Tyrou D., Martin-Ponchieu A., Ledoux-Andula N., Kouach M.,
 RA Jaquinod M., Subirana J.A., Sautiere P.;
 RT "Squid spermogenesis: molecular characterization of testis-specific
 RT pro-proteins";
 RL Biochem. J. 309:529-534 (1995).
 RN [2]
 RP PROTEIN SEQUENCE.
 RX PubMed=7626016;
 RA Moutere-Tyrou D., Martin-Ponchieu A., Ledoux-Andula N., Kouach M.,
 RA Jaquinod M., Subirana J.A., Sautiere P.;
 RT "Squid spermogenesis: molecular characterization of testis-specific
 RT pro-proteins";
 RL Biochem. J. 309 (Pt 2):529-534 (1995).
 DR PIR: S56116; S56116.
 SQ SEQUENCE 79 AA; 10788 MW; 700CBED7397E442A CRC64;
 Qy Query Match 58.7%; Score 44; DB 2; Length 79;
 Best Local Similarity 53.3%; Pred. No. 5.8;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Db 1 KGGRRMLKSTRORR 15
 19 KGGRRRRRRRRRR 33
 RESULT 9
 ID Q84RV6 ORYSA PRELIMINARY; PRT; 173 AA.
 AC Q84RV6;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein P0571D04.112.
 GN Name=P0571D04.112;
 OS Oryza sativa (Japonica cultivar-group);
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriacridae; Oryzae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa n1pobare(GA3) genomic DNA, chromosome 7, PAC
 RT clone:R0571D04";
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AP004315; BAC5921.1; -, Genomic_DNA.
 DR Grame; Q84RV6; -;
 KW Hypothetical protein.
 SQ SEQUENCE 173 AA; 18972 MW; 23B48416673AB67 CRC64;
 Qy Query Match 58.7%; Score 44; DB 2; Length 173;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Db 1 KGGRRMLKSTRORR 15
 36 KGGRRRRRRRRRR 50
 RESULT 10
 ID Q8C0D6 MOUSE PRELIMINARY; PRT; 409 AA.
 AC Q8C0D6;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
 DE enriched library, clone:603046M24 product:similar to D1223B3.1
 DE (PUTATIVE SECRETED PROTEIN ZS13), full insert sequence.
 GN Name=P78635;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=1076861; DOI=10.1101/gr.152600;
 RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bulc C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
 RA Wyshaw-Scott A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kiteuna T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasaiwa K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multiplexed sequencing";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akiyama T., Hara A., Hashizume W.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hasegawa T.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Konda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,

RA Nishi K., Nomura K., Numazeki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AK016444; BAC7491.1; -; mRNA.
 DR MGI: 2444800; P8835.
 DR MGI: 2444800; P8835.
 DR GO: 0005615; C:extracellular space; TAS.
 DR InterPro: IPR001254; Peptidase_S1_S6.
 DR Pfam: PF00089; Trypsin_1.
 DR SMART: SMO020; Tryp_Spc_1.
 KM Hydroxylase; Protease; Serine protease.
 SQ SEQUENCE 409 AA; 45736 MW; E9C9B1E6800719C CRC64;

Query Match 58.7%; Score 44; DB 2; Length 409;
 Best Local Similarity 53.3%; Pred. No. 34;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

1 KGGKRLKSTRORR 15
 194 KGGKRLKSTRORR 208

RESULT 11
 08COP9_MOUSE PRELIMINARY; PRT; 409 AA.
 AC 08COP9;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
 DE cDNA library, clone:603042412 product:similar to D02323.1
 DE (PUTATIVE SECRETED PROTEIN ZS1G13), full insert sequence (Protease,
 DE Berlin, 35).
 GN Name=Pr8835;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 ON NCBI_TaxId=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kaenaka T., Saito R.,
 RA Kaota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzei J., Mombereite P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki K., Sato K., Schoenbach C., Seta Y., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontsuki S.,
 RA Hayashizaki Y.,
 RL Functional annotation of a full-length mouse cDNA collection.",
 RN Nature 409:685-690(2001).
 RN [3]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kaenaka T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Ohsato N., Saito R., Suzuki H., Yamada I., Kiyosawa H.,
 RA Yagi K., Tomaru A., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schiraldi L.M., Kanapin A., Matuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Brad T., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragan T.A., Fletcher C.P., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guenichon S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kiyosawa A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,
 RA Maglott D.R., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.V., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Warande Y., Wells C.,
 RA Wilmink L.G., Wyszew-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavoian M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes",
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11078661; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitamura T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kasai H.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Matabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.",
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kaenaka T., Saito R.,
 RA Kaota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzei J., Mombereite P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki K., Sato K., Schoenbach C., Seta Y., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontsuki S.,
 RA Hayashizaki Y.,
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 RN Nature 409:685-690(2001).
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Eye;

```

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausener R.D., Feinold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkin R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diachenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedon T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Motley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallatton D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Matra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Beye;
RA Strauberg R.;
RL Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK011411; BAC27392.1; -; mRNA.
DR EMBL; BC075675; AAH75675.1; -; mRNA.
DR Ensemble; ENSMUSG0000033491; Mus musculus.
DR MGI; MGI:2444800; P8835.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase_S1_S6.
DR InterPro; IPR001314; Peptidase_S1A.
DR PRINTS; P00722; CHYMOTRYPSIN.
DR SMART; SM0020; TRYPSIN; UNKOWN_1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKOWN_1.
DR Hydrolase; Protease; Serine protease.
SQ SEQUENCE 409 AA; 45787 MW; 6E22DA908E7EF5 CRC64;

Query Match 58.7%; Score 44; DB 2; Length 409;
Best Local Similarity 53.3%; Fred. No. 34;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Cy 1 KGRKMLKSTRROR 15
Db 194 KGRKRLKSTRROR 208

RESULT 12
O8C0L5 MOUSE PRELIMINARY; PRT; 409 AA.
ID O8C0L5_5
AC O8C0L5_5
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched
DE library, clone:543041704 product:similar to D1223B3.1 (PUTATIVE
DE SECRETED PROTEIN ZSIC13), full insert sequence.
GN Name=P8835;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=9927953; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.

```

```

RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Komno H., Aedechi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleishmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Knehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and substructure of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Komno H., Akiyama J., Nishi K., Kiteuna T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA Aedechi J., Aizawa K., Akimura T., Hara A., Hashizume W.,
RA Fukuda S., Furuno M., Hanagata T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Komno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK030671; BAC27073.1; -; mRNA.
DR Ensemble; ENSMUSG0000033491; Mus musculus.
DR MGI; MGI:2444800; P8835.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase_S1_S6.

```


DR InterPro; IPR001314; Peptidase_S1A.
 DR PRINTS; PR000722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc_1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 KW Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 409 AA; 45847 MW; ASACICG828FD352E CRC64;

Query Match 58.7%; Score 44; DB 2; Length 409;
 Best Local Similarity 53.3%; Pred. No. 34;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KGGKRLKSTRQR 15
 ||||| : : : |||
 Db 194 KGGKRLKSTRQR 208

RESULT 13
 Q7PZL1_ANOGA PRELIMINARY; PRT; 773 AA.

AC Q7PZL1; 01-MAR-2004 (TRENBLrel. 26, Created)
 DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE ENSANGP0000008690 (Fragment).
 GN ORFNames=ENSANGS0000006546;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
 CC Anophelinae; Anopheles.
 OX NCBI_TaxID=180454;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RT "Anopheles gambiae re-annotation."
 RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RT "Anopheles gambiae Sequence Committee."
 RL Submitted (APR-2004) to the EMBL/Genbank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL; AAAB01008986; EAA00378.2; -; Genomic_DNA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR InterPro; IPR003959; AAA_Arpase_Centr.
 DR InterPro; IPR000048; IQ_Region.
 DR Pfam; PF00004; AAA; 1.
 DR Pfam; PF00612; IQ; 1.
 DR PROSITE; PS50096; IQ; 1.
 DR ATP-binding; Nucleotide-binding.
 FT NON_TER 1
 FT NON_TER 773
 SQ SEQUENCE 773 AA; 89340 MW; 5B67B1DCF48C2784 CRC64;

Query Match 58.7%; Score 44; DB 2; Length 773;
 Best Local Similarity 46.7%; Pred. No. 67;
 Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 KGGKRLKSTRQR 15
 ||||| : : : |||
 Db 427 KGGKRLKSTRQR 441

RESULT 14
 Q9WX07_STRCO PRELIMINARY; PRT; 32 AA.

DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Hypothetical protein SC03327.

GN OrderedLocustNames=SC03327; ORFNames=SC68.25C;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;

RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleiser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 RA Radniewski E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Barrall B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 DR EMBL; AL939116; CAB45361.1; -; Genomic_DNA.
 DR PIR; T36275; T36275.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 32 AA; 4045 MW; 3DE18FA26371F5A9 CRC64;

Query Match 57.3%; Score 43; DB 2; Length 32;
 Best Local Similarity 66.7%; Pred. No. 3.4;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 KGGKRLKSTRQR 15
 ||||| : : : |||
 Db 15 KGGKRLKSTRQR 29

RESULT 15
 Q4USX2_CORJK PRELIMINARY; PRT; 33 AA.
 ID Q4USX2_CORJK
 DT 13-SEP-2005 (TRENBLrel. 31, Created)
 DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=Jk1904;
 OS Corynebacterium jeikeium (strain K411).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=306537;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K411;
 RX PubMed=15968079; DOI=10.1128/JB.187.13.4671-4682.2005;

RA Tauch A., Kaiser O., Hain T., Goessmann A., Weisshar B.,
 RA Albermeier A., Beke T., Bischoff N., Brune I., Chakraborty T.,
 RA Kalinowski J., Meyer F., Rupp O., Schenker S., Viehaver P.,
 RA Puhler A.;
 RT "Complete Genome Sequence and Analysis of the Multiresistant
 RT Nosocomial Pathogen Corynebacterium jeikeium K411, a Lipid-Regulating
 RT Bacterium of the Human Skin Flora."
 RL J. Bacteriol. 187:4671-4682(2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K411;
 RX PubMed=15968079; DOI=10.1128/JB.187.13.4671-4682.2005;

RA Tauch A., Kaiser O., Hain T., Goessmann A., Weisshar B.,
 RA Albermeier A., Beke T., Bischoff N., Brune I., Chakraborty T.,
 RA Kalinowski J., Meyer F., Rupp O., Schenker S., Viehaver P.,
 RA Puhler A.;
 RT "Complete Genome Sequence and Analysis of the Multiresistant
 RT Nosocomial Pathogen Corynebacterium jeikeium K411, a Lipid-Regulating
 RT Bacterium of the Human Skin Flora."
 RL J. Bacteriol. 187:4671-4682(2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K411;
 RX PubMed=15968079; DOI=10.1128/JB.187.13.4671-4682.2005;

Query Match 57.3%; Score 43; DB 2; Length 33;
 Best Local Similarity 66.7%; Pred. No. 3.5;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy	1	KGGRKMLKSTRBORR	15
		:	
Db	15	KGGRKMLRRTIVQRR	29

Search completed: February 4, 2006, 03:41:26
Job time : 170.444 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 03:41:44 / Search time 38.3333 Seconds
(without alignments)
32.351 Million cell updates/sec

Title: US-10-790-768a-2
Perfect score: 75
Sequence: 1 KGGKMKLSTRQR 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents AA: *
1: /cgn2_6/ptodaca/1/1aa/5 COMB.pep: *
2: /cgn2_6/ptodaca/1/1aa/6 COMB.pep: *
3: /cgn2_6/ptodaca/1/1aa/H.COMB.pep: *
4: /cgn2_6/ptodaca/1/1aa/PTCUS.COMB.pep: *
5: /cgn2_6/ptodaca/1/1aa/RB.COMB.pep: *
6: /cgn2_6/ptodaca/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	57.3	33	US-09-605-703B-2766	Sequence 2766, Ap
2	42	56.0	174	US-09-107-532A-3699	Sequence 3699, Ap
3	42	56.0	248	US-09-543-681A-4404	Sequence 4404, Ap
4	42	56.0	537	US-09-252-991A-25958	Sequence 25958, A
5	41	54.7	413	US-09-999-833A-296	Sequence 296, App
6	41	54.7	413	US-10-020-445A-296	Sequence 296, App
7	41	54.7	660	US-09-252-991A-22842	Sequence 22842, A
8	41	54.7	725	US-09-252-991A-23752	Sequence 23752, A
9	40.5	54.0	204	US-09-252-991A-27153	Sequence 27153, A
10	40	53.3	160	US-09-134-001C-3904	Sequence 3904, Ap
11	40	53.3	357	US-09-538-092-254	Sequence 254, App
12	40	53.3	408	US-09-252-991A-32828	Sequence 32828, A
13	40	53.3	441	US-09-583-110-3763	Sequence 3763, Ap
14	40	53.3	442	US-09-107-433-4820	Sequence 4820, Ap
15	39	52.0	85	US-09-107-433-4454	Sequence 4454, Ap
16	39	52.0	131	US-09-513-999C-8105	Sequence 8105, Ap
17	39	52.0	277	US-09-248-796A-1606	Sequence 1606, A
18	39	52.0	1291	US-09-252-991A-19504	Sequence 19504, A
19	38	50.7	247	US-09-902-540-10732	Sequence 10732, A
20	38	50.7	305	US-09-248-588-11	Sequence 11, App1
21	38	50.7	312	US-09-134-001C-3699	Sequence 3699, Ap
22	38	50.7	404	US-09-252-991A-25830	Sequence 25830, A
23	38	50.7	494	US-09-252-991A-28162	Sequence 28162, A
24	37	49.3	44	US-09-732-210-933	Sequence 933, App
25	37	49.3	44	US-09-732-210-945	Sequence 945, App
26	37	49.3	48	US-09-732-210-948	Sequence 948, App
27	37	49.3	147	US-09-252-991A-29687	Sequence 29687, A

28	37	49.3	158	2	US-09-949-016-7013	Sequence 7013, Ap
29	37	49.3	173	2	US-09-107-532A-4569	Sequence 4569, Ap
30	37	49.3	174	1	US-08-933-750C-27	Sequence 27, Appl
31	37	49.3	174	2	US-09-234-613-27	Sequence 27, Appl
32	37	49.3	174	2	US-09-949-016-6589	Sequence 6589, Ap
33	37	49.3	187	2	US-09-949-016-8739	Sequence 8739, Ap
34	37	49.3	237	2	US-09-248-796A-20114	Sequence 20114, A
35	37	49.3	248	2	US-09-252-991A-32920	Sequence 32920, A
36	37	49.3	299	2	US-08-879-098-2	Sequence 2, Appli
37	37	49.3	299	2	US-09-631-548-2	Sequence 2, Appli
38	37	49.3	309	2	US-09-252-991A-27914	Sequence 27914, A
39	37	49.3	313	2	US-09-252-991A-28257	Sequence 28257, A
40	37	49.3	345	2	US-09-134-000C-5685	Sequence 5685, Ap
41	37	49.3	351	2	US-09-107-433-4511	Sequence 4511, Ap
42	37	49.3	575	2	US-09-252-991A-25723	Sequence 25723, A
43	37	49.3	639	2	US-09-252-991A-28453	Sequence 28453, A
44	37	49.3	704	2	US-09-252-991A-17523	Sequence 17523, A
45	37	49.3	738	2	US-08-864-038A-3	Sequence 3, Appl1

ALIGNMENTS

```

RESULT 1
US-09-605-703B-2766
: Sequence 2766, Application US/09605703B
: Patent No. 6962989
: GENERAL INFORMATION:
: APPLICANT: Pompeius, Markus
: APPLICANT: Krogger, Burkhard
: APPLICANT: Schröder, Hartwig
: APPLICANT: Zeidler, Oskar
: APPLICANT: Haberhauser, Gregor
: TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
: FILE REFERENCE: BGI-129CP
: CURRENT APPLICATION NUMBER: US/09/605, 703B
: PRIOR FILING DATE: 2000-06-27
: PRIOR APPLICATION NUMBER: 60/142,764
: PRIOR FILING DATE: 1999-07-08
: PRIOR APPLICATION NUMBER: 60/152,318
: NUMBER OF SEQ ID NOS: 2934
: SEQ ID NO 2766
: LENGTH: 33
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
US-09-605-703B-2766

Query Match      57.3%; Score 43; DB 2; Length 33;
Best Local Similarity 66.7%; Pred. No. 2.3;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy      1  KGGKMKLSTRQR 15
Db      15  KGGKMKLSTRQR 29

RESULT 2
US-09-107-532A-3699
: Sequence 3699, Application US/09107532A
: Patent No. 6583275
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Walham
STATE: Massachusetts
COUNTRY: USA

```

```
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3699:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1..174
SEQUENCE DESCRIPTION: SEQ ID NO: 3699:
US-09-107-532A-3699

Query Match          56.0%; Score 42; DB 2; Length 174;
Best Local Similarity 57.1%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 GGRKMLKSTRROR 15
       |||||:|:|
Db      7 GGRKMKIKAVRIDER 20

RESULT 3
US-09-543-681A-4404
Sequence 4404, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709,1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 4404
LENGTH: 248
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-4404

Query Match          56.0%; Score 42; DB 2; Length 248;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      2 GGRKMLKSTRROR 14
       |||||:|:|
Db     139 GGRRLATTLROR 151
```

```
RESULT 4
US-09-252-991A-25958
Sequence 25958, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25958
LENGTH: 537
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25958

Query Match          56.0%; Score 42; DB 2; Length 537;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      2 GGRKMLKSTRROR 15
       |||||:|:|
Db     400 GSRRLKTRRRR 413

RESULT 5
US-09-999-833A-296
Sequence 296, Application US/09999833A
Patent No. 6916648
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Deenoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Fliviaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottlisen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C65
CURRENT APPLICATION NUMBER: US/09/999,833A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
```

1	PRIOR APPLICATION NUMBER: 1997-11-03	60/064243
2	PRIOR FILING DATE: 1997-11-03	
3	PRIOR APPLICATION NUMBER: 60/065311	11-13
4	PRIOR FILING DATE: 1997-11-13	
5	PRIOR APPLICATION NUMBER: 60/066364	66-13
6	PRIOR FILING DATE: 1997-11-21	
7	PRIOR APPLICATION NUMBER: 60/077456	
8	PRIOR FILING DATE: 1998-03-10	
9	PRIOR APPLICATION NUMBER: 60/077632	10-11
10	PRIOR FILING DATE: 1998-03-11	
11	PRIOR APPLICATION NUMBER: 60/077641	11-11
12	PRIOR FILING DATE: 1998-03-11	
13	PRIOR APPLICATION NUMBER: 60/077649	
14	PRIOR FILING DATE: 1998-03-11	
15	PRIOR APPLICATION NUMBER: 60/07773D	
16	PRIOR FILING DATE: 1998-03-12	
17	PRIOR APPLICATION NUMBER: 60/078004	
18	PRIOR FILING DATE: 1998-03-13	
19	PRIOR APPLICATION NUMBER: 60/078866	03-20
20	PRIOR FILING DATE: 1998-03-20	
21	PRIOR APPLICATION NUMBER: 60/078936	
22	PRIOR FILING DATE: 1998-03-20	
23	PRIOR APPLICATION NUMBER: 60/078910	10-20
24	PRIOR FILING DATE: 1998-03-20	
25	PRIOR APPLICATION NUMBER: 60/078939	
26	PRIOR FILING DATE: 1998-03-20	
27	PRIOR APPLICATION NUMBER: 60/079234	
28	PRIOR FILING DATE: 1998-03-25	
29	PRIOR APPLICATION NUMBER: 60/079656	03-26
30	PRIOR FILING DATE: 1998-03-26	
31	PRIOR APPLICATION NUMBER: 60/079664	03-27
32	PRIOR FILING DATE: 1998-03-27	
33	PRIOR APPLICATION NUMBER: 60/079669	
34	PRIOR FILING DATE: 1998-03-27	
35	PRIOR APPLICATION NUMBER: 60/079663	03-31
36	PRIOR FILING DATE: 1998-03-27	
37	PRIOR APPLICATION NUMBER: 60/079728	03-31
38	PRIOR FILING DATE: 1998-03-27	
39	PRIOR APPLICATION NUMBER: 60/079786	03-27
40	PRIOR FILING DATE: 1998-03-27	
41	PRIOR APPLICATION NUMBER: 60/079920	03-30
42	PRIOR FILING DATE: 1998-03-30	
43	PRIOR APPLICATION NUMBER: 60/079923	03-30
44	PRIOR FILING DATE: 1998-03-30	
45	PRIOR APPLICATION NUMBER: 60/080105	03-31
46	PRIOR FILING DATE: 1998-03-31	
47	PRIOR APPLICATION NUMBER: 60/080107	
48	PRIOR FILING DATE: 1998-03-31	
49	PRIOR APPLICATION NUMBER: 60/080165	
50	PRIOR FILING DATE: 1998-03-31	
51	PRIOR APPLICATION NUMBER: 60/080194	
52	PRIOR FILING DATE: 1998-03-31	
53	PRIOR APPLICATION NUMBER: 60/080327	
54	PRIOR FILING DATE: 1998-04-01	
55	PRIOR APPLICATION NUMBER: 60/080328	
56	PRIOR FILING DATE: 1998-04-01	
57	PRIOR APPLICATION NUMBER: 60/080333	04-01
58	PRIOR FILING DATE: 1998-04-01	
59	PRIOR APPLICATION NUMBER: 60/080334	04-01
60	PRIOR FILING DATE: 1998-04-01	
61	PRIOR APPLICATION NUMBER: 60/081070	04-08
62	PRIOR FILING DATE: 1998-04-08	
63	PRIOR APPLICATION NUMBER: 60/081049	04-08
64	PRIOR FILING DATE: 1998-04-08	
65	PRIOR APPLICATION NUMBER: 60/081071	11-11
66	PRIOR FILING DATE: 1998-04-08	
67	PRIOR APPLICATION NUMBER: 60/081195	
68	PRIOR FILING DATE: 1998-04-08	
69	PRIOR APPLICATION NUMBER: 60/081203	04-09
70	PRIOR FILING DATE: 1998-04-09	
71	PRIOR APPLICATION NUMBER: 60/081229	
72	PRIOR FILING DATE: 1998-04-09	
73	PRIOR APPLICATION NUMBER: 60/081955	

PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/0081817
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/0081819
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/0081955
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/0081838
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/0082568
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION NUMBER:	60/0082565
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION NUMBER:	60/0082704
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/0082804
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/0082700
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/0082797
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/0082796
PRIOR FILING DATE:	1998-04-23
PRIOR APPLICATION NUMBER:	60/0083336
PRIOR FILING DATE:	1998-04-27
PRIOR APPLICATION NUMBER:	60/0083322
PRIOR FILING DATE:	1998-04-28
PRIOR APPLICATION NUMBER:	60/0083392
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083499
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083455
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083495
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083466
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083558
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083500
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083430
PRIOR FILING DATE:	1998-05-05
PRIOR APPLICATION NUMBER:	60/0084414
PRIOR FILING DATE:	1998-05-06
PRIOR APPLICATION NUMBER:	60/0084441
PRIOR FILING DATE:	1998-05-06
PRIOR APPLICATION NUMBER:	60/0084637
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0084639
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0084640
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0084598
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0084600
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0084627
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0084633
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/0085338
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/0085323
PRIOR FILING DATE:	1998-05-13

PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 54.7%; Score 41; DB 2; Length 413;
Best Local Similarity 64.3%; Pred. No. 53;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 KGRKMLKSTRROR 14
|||: |||
Db 228 KGRKRRKSGRGR 241

RESULT 6
US-10-020-445A-296
Sequence 296, Application US/10020445A
Patent No. 6962797
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Batoni, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottlieb, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C74
CURRENT APPLICATION NUMBER: US/10/020,445A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450

PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15

PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579

PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 54.7%; Score 41; DB 2; Length 413;
Best Local Similarity 64.3%; Pred. No. 53;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KGGKMKSTRROR 14
Db 228 KGGRRKKSGRGOR 241

RESULT 7
US-09-252-991A-22842
Sequence 22842, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22842
LENGTH: 660
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22842

Query Match 54.7%; Score 41; DB 2; Length 660;
Best Local Similarity 61.5%; Pred. No. 83;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 3 GRKMKSTRROR 15
Db 602 GRRLPQARROR 614

RESULT 8
US-09-252-991A-23752
Sequence 23752, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23752
LENGTH: 725
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23752

Query Match 54.7%; Score 41; DB 2; Length 725;
Best Local Similarity 61.5%; Pred. No. 90;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 3 GGRKMLKSTRRQR 15
 Db 646 GRRQLPQARRRR 658

RESULT 9
 US-09-252-991A-27153
 ; Sequence 27153, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 27153
 ; LENGTH: 204
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-27153

Query Match 54.0%; Score 40.5; DB 2; Length 204;
 Best Local Similarity 56.2%; Pred. No. 33;
 Matches 9; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 1 KGRK-MLKSTRRQR 15
 Db 184 RGRRRVARSTRRAR 199

RESULT 10
 US-09-134-001C-3904
 ; Sequence 3904, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 3904
 ; LENGTH: 160
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-3904

Query Match 53.3%; Score 40; DB 2; Length 160;
 Best Local Similarity 53.8%; Pred. No. 32;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KGRKMLKSTRRQR 13
 Db 3 RGRTRVKPTRRQR 15

RESULT 11
 US-09-538-092-254
 ; Sequence 254, Application US/09538092
 ; Patent No. 6753314

; GENERAL INFORMATION:
 ; APPLICANT: Glot, Loic
 ; APPLICANT: Mansfield, Traci A.
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
 ; FILE REFERENCE: 15966-542
 ; CURRENT APPLICATION NUMBER: US/09/538,092
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/127,352
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178,965
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 1387
 ; SOFTWARE: CurePatSeqFormatter Version 0.9
 ; SEQ ID NO 254
 ; LENGTH: 357
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (0)...(0)
 ; OTHER INFORMATION: Polypeptide Accession Number YER127W
 US-09-538-092-254

Query Match 53.3%; Score 40; DB 2; Length 357;
 Best Local Similarity 46.7%; Pred. No. 67;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 KGRKMLKSTRRQR 15
 Db 271 RGRKSLRDEKRR 285

RESULT 12
 US-09-252-991A-32828
 ; Sequence 32828, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 32828
 ; LENGTH: 408
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-32828

Query Match 53.3%; Score 40; DB 2; Length 408;
 Best Local Similarity 50.0%; Pred. No. 76;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGRKMLKSTRRQR 15
 Db 87 GRRQLRRRRRRR 100

RESULT 13
 US-09-583-110-3763
 ; Sequence 3763, Application US/09583110
 ; Patent No. 6699703
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al.
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
 ; FILE REFERENCE: PATH00-07A
 ; CURRENT APPLICATION NUMBER: US/09/583,110

CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3763
LENGTH: 441
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3763

Query Match 53.3%; Score 40; DB 2; Length 441;
Best Local Similarity 53.8%; Pred. No. 82;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KGRKMLKSTRQ 13
||:||||:
Db 416 KGRKMLKASKTQ 428

RESULT 14
US-09-107-433-4820
Sequence 4820, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4820:
SEQUENCE CHARACTERISTICS:
LENGTH: 442 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...442
SEQUENCE DESCRIPTION: SEQ ID NO: 4820:

US-09-107-433-4820

Query Match 53.3%; Score 40; DB 2; Length 442;
Best Local Similarity 53.8%; Pred. No. 82;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KGRKMLKSTRQ 13
||:||||:
Db 417 KGRKMLKASKTQ 429

RESULT 15
US-09-107-433-4454
Sequence 4454, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4454:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...85
SEQUENCE DESCRIPTION: SEQ ID NO: 4454:
US-09-107-433-4454

Query Match 52.0%; Score 39; DB 2; Length 85;
Best Local Similarity 53.8%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KGRKMLKSTRQ 13
||:||||:
Db 3 KGNQMVVKTQ 15

Search completed: February 4, 2006, 03:43:34

Mon Feb 6 10:11:05 2006

us-10-790-768a-2.ra1

Page 8

Job time : 39.333 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 03:42:34 : Search time 130 Seconds
(without alignments)
48.211 Million cell updates/sec

Title: US-10-790-768A-2

Sequence: 1 KGRKMLKSTRQR 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing filter: 45 summaries

Database:

1: Published Applications AA Main: *
2: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep: *
3: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep: *
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep: *
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep: *
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	15	4 US-10-790-768A-2	Sequence 2, Appl1
2	75	100.0	15	4 US-10-790-768A-14	Sequence 14, Appl1
3	75	100.0	21	4 US-10-790-768A-4	Sequence 4, Appl1
4	75	100.0	21	4 US-10-790-768A-19	Sequence 19, Appl1
5	75	100.0	21	4 US-10-790-768A-20	Sequence 20, Appl1
6	70	93.3	15	4 US-10-790-768A-16	Sequence 16, Appl1
7	62	82.7	27	4 US-10-790-768A-5	Sequence 5, Appl1
8	60	80.0	19	4 US-10-790-768A-18	Sequence 18, Appl1
9	60	80.0	25	4 US-10-790-768A-21	Sequence 21, Appl1
10	60	80.0	25	4 US-10-790-768A-22	Sequence 22, Appl1
11	58	77.3	12	4 US-10-790-768A-1	Sequence 1, Appl1
12	48	64.0	123	4 US-10-437-963-117103	Sequence 117103,
13	45	60.0	71	4 US-10-425-115-335780	Sequence 335780,
14	45	60.0	138	4 US-10-437-963-188683	Sequence 188683,
15	43	57.3	33	3 US-09-738-626-3964	Sequence 3964, Ap
16	43	57.3	247	4 US-10-408-765A-2219	Sequence 2219, Ap
17	43	57.3	282	4 US-10-437-963-122500	Sequence 122500,
18	42	56.0	145	4 US-10-424-599-176649	Sequence 176649,
19	42	56.0	708	4 US-10-408-765A-1772	Sequence 1772, Ap
20	41	54.7	153	4 US-10-425-115-225169	Sequence 225169,
21	41	54.7	155	4 US-10-425-115-347520	Sequence 347520,
22	41	54.7	413	3 US-09-888-615-89	Sequence 89, Appl
23	41	54.7	413	3 US-09-978-295A-296	Sequence 296, App
24	41	54.7	413	3 US-09-978-697-296	Sequence 296, App
25	41	54.7	413	3 US-09-978-192A-296	Sequence 296, App
26	41	54.7	413	3 US-09-999-832A-296	Sequence 296, App
27	41	54.7	413	3 US-09-978-189-296	Sequence 296, App

28	41	54.7	413	3 US-09-978-608A-296	Sequence 296, App
29	41	54.7	413	3 US-09-978-585A-296	Sequence 296, App
30	41	54.7	413	3 US-09-978-191A-296	Sequence 296, App
31	41	54.7	413	3 US-09-978-403A-296	Sequence 296, App
32	41	54.7	413	3 US-09-978-564A-296	Sequence 296, App
33	41	54.7	413	3 US-09-999-833A-296	Sequence 296, App
34	41	54.7	413	3 US-09-981-915A-296	Sequence 296, App
35	41	54.7	413	3 US-09-978-824-296	Sequence 296, App
36	41	54.7	413	3 US-09-918-585A-296	Sequence 296, App
37	41	54.7	413	3 US-09-999-834A-296	Sequence 296, App
38	41	54.7	413	3 US-09-978-423A-296	Sequence 296, App
39	41	54.7	413	3 US-09-978-193A-296	Sequence 296, App
40	41	54.7	413	3 US-09-999-830A-296	Sequence 296, App
41	41	54.7	413	3 US-09-978-157A-296	Sequence 296, App
42	41	54.7	413	3 US-09-978-187B-296	Sequence 296, App
43	41	54.7	413	3 US-09-978-643A-296	Sequence 296, App
44	41	54.7	413	3 US-09-978-175A-296	Sequence 296, App
45	41	54.7	413	3 US-09-978-298A-296	Sequence 296, App

ALIGNMENTS

```
RESULT 1
US-10-790-768A-2
; Sequence 2, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
US-10-790-768A-2

Query Match      100.0%; Score 75; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 KGRKMLKSTRQR 15
      |||||
Db      1 KGRKMLKSTRQR 15

RESULT 2
US-10-790-768A-14
; Sequence 14, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
```

; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768A-14

Query Match 100.0%; Score 75; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGRKMLKSTRRQR 15
Db 1 KGRKMLKSTRRQR 15

RESULT 3
US-10-790-768A-4
; Sequence 4, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
US-10-790-768A-4

Query Match 100.0%; Score 75; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGRKMLKSTRRQR 15
Db 1 KGRKMLKSTRRQR 15

RESULT 4
US-10-790-768A-19
; Sequence 19, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768A-19

Query Match 100.0%; Score 75; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGRKMLKSTRRQR 15
Db 1 KGRKMLKSTRRQR 15

RESULT 5
US-10-790-768A-20
; Sequence 20, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768A-20

Query Match 100.0%; Score 75; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGRKMLKSTRRQR 15
Db 7 KGRKMLKSTRRQR 21

RESULT 6
US-10-790-768A-16
; Sequence 16, Application US/10790768A
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790,768A
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768A-16

Query Match 93.3%; Score 70; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 9.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGRKMLKSTRRQR 15
Db 2 GGRKMLKSTRRQR 15

```
RESULT 7
US-10-790-768a-5
; Sequence 5, Application US/10790768a
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790.768a
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
US-10-790-768a-5

Query Match      82.7%; Score 62; DB 4; Length 27;
Best Local Similarity 71.4%; Pred. No. 0.0037;
Matches 15; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy      1 KGGK-----RKMKSTRROR 15
Db      1 KGGKKKKKKVLMKSTRROR 21

RESULT 8
US-10-790-768a-18
; Sequence 18, Application US/10790768a
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790.768a
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 18
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768a-18

Query Match      80.0%; Score 60; DB 4; Length 19;
Best Local Similarity 73.7%; Pred. No. 0.0056;
Matches 14; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

Qy      1 KGGRK-----MLKSTRROR 15
Db      1 KGGKKKKKKVLMKSTRROR 19

RESULT 9
US-10-790-768a-21
; Sequence 21, Application US/10790768a
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
```

```
; CURRENT APPLICATION NUMBER: US/10/790.768a
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 21
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768a-21

Query Match      80.0%; Score 60; DB 4; Length 25;
Best Local Similarity 73.7%; Pred. No. 0.0074;
Matches 14; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

Qy      1 KGGRK-----MLKSTRROR 15
Db      1 KGGKKKKKKVLMKSTRROR 19

RESULT 10
US-10-790-768a-22
; Sequence 22, Application US/10790768a
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790.768a
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 22
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protein transduction domain
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotin
US-10-790-768a-22

Query Match      80.0%; Score 60; DB 4; Length 25;
Best Local Similarity 73.7%; Pred. No. 0.0074;
Matches 14; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

Qy      1 KGGRK-----MLKSTRROR 15
Db      7 KGGKKKKKKVLMKSTRROR 25

RESULT 11
US-10-790-768a-1
; Sequence 1, Application US/10790768a
; Publication No. US20040209797A1
; GENERAL INFORMATION:
; APPLICANT: Karas, Michael
; TITLE OF INVENTION: Intracellular Delivery of Small Molecules, Proteins, and Nucleic
; FILE REFERENCE: 002877.00028
; CURRENT APPLICATION NUMBER: US/10/790.768a
; CURRENT FILING DATE: 2004-03-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
```

LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: protein transduction domain
US-10-790-768a-1

Query Match 77.3%; Score 58; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RKMLKSTRRQR 15
Db 1 RKMLKSTRRQR 12

RESULT 12
US-10-437-963-117103
Sequence 117103, Application US/10437963
Publication No. US2004012343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovall, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 117103
LENGTH: 123
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_20540C.1.pep
US-10-437-963-117103

Query Match 64.0%; Score 48; DB 4; Length 123;
Best Local Similarity 69.2%; Pred. No. 3.8;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KGRKMLKSTRRQ 13
Db 11 KGRKMLKSTRRQ 23

RESULT 13
US-10-425-115-335780
Sequence 335780, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovall, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 335780
LENGTH: 71
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_6933C.1.pep

US-10-425-115-335780

Query Match 60.0%; Score 45; DB 4; Length 71;
Best Local Similarity 64.3%; Pred. No. 6.8;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGRKMLKSTRRQR 15
Db 12 GGRKMLKSTRRQR 25

RESULT 14
US-10-437-963-188683
Sequence 188683, Application US/10437963
Publication No. US2004012343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovall, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 188683
LENGTH: 128
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(128)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_85263C.1.pep
US-10-437-963-188683

Query Match 60.0%; Score 45; DB 4; Length 128;
Best Local Similarity 64.3%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGRKMLKSTRRQR 15
Db 112 GGRKMLKSTRRQR 125

RESULT 15
US-09-738-626-3964
Sequence 3964, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENO, AKIHIRO
APPLICANT: IKEDA, MASATO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162

/ PRIOR FILING DATE: 2000-04-07
 / PRIOR APPLICATION NUMBER: JP 00/280988
 / PRIOR FILING DATE: 2000-08-03
 / NUMBER OF SEQ ID NOS: 7059
 / SOFTWARE: PatentIn ver. 3.0
 / SEQ ID NO 3964
 / LENGTH: 33
 / TYPE: PRT
 / ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3964

Query Match 57.3%; Score 43; DB 3; Length 33;
 Best Local Similarity 66.7%; Pred. No. 6.7;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KGGKMKLSTRRQR 15
 DB 15 KKHKKMLRTRVQR 29

Search completed: February 4, 2006, 03:47:34
 Job time : 131 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 03:43:49 ; Search time 10.5556 Seconds
(without alignments)
16.652 Million cell updates/sec

Title: US-10-790-768a-2

Perfect score: 75
Sequence: 1 KGRKMLKSTRRQR 15

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	41	54.7	413	US-11-037-243-89	Sequence 89, Appl
2	38	50.7	672	US-10-689-742-70	Sequence 70, Appl
3	37	49.3	54	US-11-150-054A-13	Sequence 13, Appl
4	37	49.3	54	US-11-150-054A-21	Sequence 21, Appl
5	37	49.3	54	US-11-150-054A-26	Sequence 26, Appl
6	36	48.0	154	US-10-821-234-1344	Sequence 1344, Ap
7	36	48.0	280	US-10-821-234-1300	Sequence 1300, Ap
8	35.5	47.3	327	US-11-165-211-52	Sequence 52, Appl
9	35.5	47.3	327	US-11-165-226-62	Sequence 62, Appl
10	35	46.7	342	US-10-714-887-10	Sequence 10, Appl
11	35	46.7	441	US-11-024-959-410	Sequence 410, Appl
12	35	46.7	844	US-10-453-372-852	Sequence 852, Appl
13	35	46.7	844	US-10-453-372-856	Sequence 856, Appl
14	34	45.3	356	US-10-161-408-50	Sequence 50, Appl
15	34	45.3	543	US-11-169-041-162	Sequence 162, Appl
16	33	44.0	441	US-11-024-959-280	Sequence 280, Appl
17	33	44.0	441	US-11-024-959-281	Sequence 281, Appl
18	33	44.0	521	US-11-091-914-2	Sequence 2, Appl1
19	33	44.0	663	US-11-024-959-314	Sequence 314, Appl
20	33	44.0	860	US-11-019-711-59	Sequence 59, Appl
21	32	42.7	99	US-10-467-657-5934	Sequence 5934, Appl
22	32	42.7	164	US-10-467-657-170	Sequence 170, Appl
23	32	42.7	164	US-10-467-657-204	Sequence 204, Appl
24	32	42.7	164	US-10-467-657-3632	Sequence 3632, Appl
25	32	42.7	164	US-10-467-657-6390	Sequence 6390, Appl

26	32	42.7	178	6	US-10-467-657-8516	Sequence 8516, Appl
27	32	42.7	181	6	US-10-793-626-560	Sequence 560, Appl
28	32	42.7	211	7	US-11-214-199-10	Sequence 10, Appl1
29	32	42.7	212	7	US-11-214-199-4	Sequence 4, Appl1
30	32	42.7	212	7	US-11-214-199-12	Sequence 12, Appl1
31	32	42.7	411	7	US-11-092-168-8	Sequence 8, Appl1
32	32	42.7	718	7	US-11-024-959-273	Sequence 273, Appl
33	32	42.7	784	7	US-11-147-109-4	Sequence 4, Appl1
34	32	42.7	874	7	US-11-012-762-8	Sequence 8, Appl1
35	32	42.7	882	7	US-11-012-762-34	Sequence 34, Appl1
36	31	41.3	27	7	US-11-096-706-214	Sequence 214, Appl
37	31	41.3	27	7	US-11-101-287-159	Sequence 159, Appl
38	31	41.3	27	7	US-11-101-287-219	Sequence 219, Appl
39	31	41.3	27	7	US-11-078-469-32	Sequence 32, Appl1
40	31	41.3	27	7	US-11-225-686-11	Sequence 11, Appl1
41	31	41.3	28	7	US-11-078-469-33	Sequence 33, Appl1
42	31	41.3	30	7	US-11-078-469-24	Sequence 24, Appl1
43	31	41.3	30	7	US-11-078-469-62	Sequence 62, Appl1
44	31	41.3	119	7	US-11-120-308-146	Sequence 146, Appl
45	31	41.3	126	6	US-10-467-657-892	Sequence 892, Appl

ALIGNMENTS

RESULT 1
US-11-037-243-89
; Sequence 89, Application US/11037243
; Publication No. US20050287546A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CANBERRELL, SEAN
; APPLICANT: CHARVOTZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/11/037,243
; CURRENT FILING DATE: 2005-05-26
; PRIOR APPLICATION NUMBER: US/09/888,615
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-037-243-89
Query Match 54.7%; Score 41; DB 7; Length 413;
Best Local Similarity 64.3%; Pred. No. 5.2;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 KGRKMLKSTRRQR 14
Db 228 KGRRRKSGRGQR 241
RESULT 2
US-10-689-742-70
; Sequence 70, Application US/10689742
; Publication No. US20050250180A1
; GENERAL INFORMATION:
; APPLICANT: JACOBS, KENNETH
; APPLICANT: MCCOY, JOHN M
; APPLICANT: LAVAILLE, EDWARD R
; APPLICANT: RACIE, LISA A
; APPLICANT: EVANS, CHERYL
; APPLICANT: MERBERG, DAVID
; APPLICANT: TREACY, MAURICE

```
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 00766,000091.10
CURRENT APPLICATION NUMBER: US/10/689,742
CURRENT FILING DATE: 2003-10-22
PRIOR APPLICATION NUMBER: 09/746,783
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 231
SOFTWARE: PatentIn version 3.2
SEQ ID NO 70
LENGTH: 672
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (107)..(107)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (111)..(111)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (117)..(118)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (645)..(645)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-689-742-70
```

```
Query Match          50.7% Score 38; DB 6; Length 672;
Best Local Similarity 53.8%; Pred. No. 28;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 3 GRKMLKSTRRQR 15
|||::|||
Db 597 GRKLEAGRRESR 609
```

```
RESULT 3
US-11-150-054A-13
Sequence 13, Application US/11150054A
Publication No. US20050278801A1
GENERAL INFORMATION:
APPLICANT: Laessner, Michael Q.
TITLE OF INVENTION: Plasticid Transit Peptides
FILE REFERENCE: 2119-4284US1
CURRENT APPLICATION NUMBER: US/11/150,054A
CURRENT FILING DATE: 2005-06-09
PRIOR APPLICATION NUMBER: 60/578,535
PRIOR FILING DATE: 2004-06-09
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn version 3.3
SEQ ID NO 13
LENGTH: 54
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-11-150-054A-13
```

```
Query Match          49.3% Score 37; DB 7; Length 54;
Best Local Similarity 58.3%; Pred. No. 3.3;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 4 RKMLKSTRRQR 15
|:::|||||
Db 34 RRFNNTRRQR 45
```

RESULT 4

```
US-11-150-054A-21
Sequence 21, Application US/11150054A
Publication No. US20050278801A1
GENERAL INFORMATION:
APPLICANT: Laessner, Michael Q.
TITLE OF INVENTION: Plasticid Transit Peptides
FILE REFERENCE: 2119-4284US1
CURRENT APPLICATION NUMBER: US/11/150,054A
CURRENT FILING DATE: 2005-06-09
PRIOR APPLICATION NUMBER: 60/578,535
PRIOR FILING DATE: 2004-06-09
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn version 3.3
SEQ ID NO 21
LENGTH: 54
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-11-150-054A-21
```

```
Query Match          49.3% Score 37; DB 7; Length 54;
Best Local Similarity 58.3%; Pred. No. 3.3;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 4 RKMLKSTRRQR 15
|:::|||||
Db 34 RRFNNTRRQR 45
```

```
RESULT 5
US-11-150-054A-26
Sequence 26, Application US/11150054A
Publication No. US20050278801A1
GENERAL INFORMATION:
APPLICANT: Laessner, Michael Q.
TITLE OF INVENTION: Plasticid Transit Peptides
FILE REFERENCE: 2119-4284US1
CURRENT APPLICATION NUMBER: US/11/150,054A
CURRENT FILING DATE: 2005-06-09
PRIOR APPLICATION NUMBER: 60/578,535
PRIOR FILING DATE: 2004-06-09
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn version 3.3
SEQ ID NO 26
LENGTH: 54
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-11-150-054A-26
```

```
Query Match          49.3% Score 37; DB 7; Length 54;
Best Local Similarity 58.3%; Pred. No. 3.3;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 4 RKMLKSTRRQR 15
|:::|||||
Db 34 RRFNNTRRQR 45
```

```
RESULT 6
US-10-821-234-1344
Sequence 1344, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarment, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
```

FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821.234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PC_SEQ_genes Version 1.0
SEQ ID NO 1344
LENGTH: 154
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1344

Query Match 48.0%; Score 36; DB 6; Length 154;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGRKMLKSTRRQR 15
DB 72 GGRDLESSGLQR 85

RESULT 7
US-10-821-234-1300
Sequence 1300, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crahn, Birgit
APPLICANT: Andarman, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821.234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PC_SEQ_genes Version 1.0
SEQ ID NO 1300
LENGTH: 280
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1300

Query Match 48.0%; Score 36; DB 6; Length 280;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGRKMLKSTRRQR 15
DB 20 GDRGLVSTRREER 33

RESULT 8
US-11-165-211-52
Sequence 52, Application US/11165211
Publication No. US20050287626A1
GENERAL INFORMATION:
APPLICANT: KYOWA HAKKO KOGYO CO., LTD.
TITLE OF INVENTION: Process for producing dipeptides
FILE REFERENCE: 4093-14
CURRENT APPLICATION NUMBER: US/11/165.211
CURRENT FILING DATE: 2005-06-24
PRIOR APPLICATION NUMBER: JP2004-189012
PRIOR FILING DATE: 2004-06-25
NUMBER OF SEQ ID NOS: 132
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 52
LENGTH: 327
TYPE: PRT
ORGANISM: Escherichia coli
US-11-165-211-52

Query Match 47.3%; Score 35.5; DB 7; Length 327;
Best Local Similarity 64.3%; Pred. No. 37;
Matches 9; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 KGRKMLKSTRRQR 14
DB 120 QGNN---KSTRRQR 130

RESULT 9
US-11-165-226-62
Sequence 62, Application US/11165226
Publication No. US20050287627A1
GENERAL INFORMATION:
APPLICANT: KYOWA HAKKO KOGYO CO., LTD.
TITLE OF INVENTION: Process for producing dipeptides or dipeptide derivatives
FILE REFERENCE: 4093-13
CURRENT APPLICATION NUMBER: US/11/165.226
CURRENT FILING DATE: 2005-06-24
PRIOR APPLICATION NUMBER: JP2004-189007
PRIOR FILING DATE: 2004-06-25
NUMBER OF SEQ ID NOS: 131
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 62
LENGTH: 327
TYPE: PRT
ORGANISM: Escherichia coli
US-11-165-226-62

Query Match 47.3%; Score 35.5; DB 7; Length 327;
Best Local Similarity 64.3%; Pred. No. 37;
Matches 9; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 KGRKMLKSTRRQR 14
DB 120 QGNN---KSTRRQR 130

RESULT 10
US-10-714-887-10
Sequence 10, Application US/10714887
Publication No. US20060015972A1
GENERAL INFORMATION:
APPLICANT: Menzel Biotechnology, Inc.
APPLICANT: HEARD, Jacqueline
APPLICANT: RIECHMANN, Jose Luis
APPLICANT: CREELMAN, Robert
APPLICANT: RATCLIFFE, Oliver
APPLICANT: CANALES, Roger
APPLICANT: REPETTI, Peter
APPLICANT: KUMIMOTO, Roderick W
APPLICANT: GUTTERSON, Neal
APPLICANT: REUBER, T. Lynne
APPLICANT: PINEDA, Omaira
APPLICANT: SHERMAN, Bradley K
TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
FILE REFERENCE: MBI0058-CIP
CURRENT APPLICATION NUMBER: US/10/714.887
CURRENT FILING DATE: 2003-11-13
PRIOR APPLICATION NUMBER: 10/412,699
PRIOR FILING DATE: 2003-04-10
PRIOR APPLICATION NUMBER: 09/506,720
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/135,134
PRIOR FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: 09/394,519
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: 09/533,392
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 09/533,029
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 09/532,591

```
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 10
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G2053 polypeptide reference sequence; clade identifier
US-10-714-887-10

Query Match          46.7%; Score 35; DB 6; Length 342;
Best Local Similarity 42.9%; Pred. No. 47;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      2 GGRRLKSTRRQR 15
      |||:|:|:|:|
DB      305 GGKLGQETREKR 318

RESULT 11
US-11-024-959-410
; Sequence 410, Application US/11024959
; Publication No. US20060010516a1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSTIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 04463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; PRIOR FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 410
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-410

Query Match          46.7%; Score 35; DB 7; Length 441;
Best Local Similarity 60.0%; Pred. No. 62;
Matches 9; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY      1 KGRRLKSTRRQR 15
      |||:|:|:|:|
DB      124 KGG--SLKMTTRHQR 136

RESULT 12
US-10-453-372-852
; Sequence 852, Application US/10453372
; Publication No. US2006000323a1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; PRIOR APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
```

```
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO: 852
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-852

Query Match          46.7%; Score 35; DB 6; Length 844;
Best Local Similarity 53.8%; Pred. No. 1,2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 GGRRLKSTRRQR 14
      |||:|:|:|:|
DB      825 GGRQLHSRRKQK 837

RESULT 13
US-10-453-372-856
; Sequence 856, Application US/10453372
; Publication No. US2006000323a1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO: 856
; LENGTH: 844
```

TYPE: PRT
ORGANISM: Homo sapiens
US-10-453-372-856

Query Match 46.7%; Score 35; DB 6; Length 844;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGRKMLKSTRROR 14
DB 825 GGRGQLHSRKROK 837

RESULT 14

US-10-161-408-50
Sequence 50, Application US/10161408
Publication No. US20050287647A9
GENERAL INFORMATION:
APPLICANT: Perez, Carl
APPLICANT: Fabijanski, Steven
APPLICANT: Perkins, Edward
TITLE OF INVENTION: Plant Artificial Chromosomes, Uses thereof, and Methods of Prepara
FILE REFERENCE: 24601-419
CURRENT APPLICATION NUMBER: US/10/161,408
PENDING FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: US 60/294,687
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: US 60/296,329
PRIOR FILING DATE: 2001-06-04
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 50
LENGTH: 356
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Integrase E174R
US-10-161-408-50

Query Match 45.3%; Score 34; DB 6; Length 356;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGRKMLKSTRROR 13
DB 267 GGRITIASTRRE 278

RESULT 15

US-11-169-041-162
Sequence 162, Application US/11169041
Publication No. US20060019284A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
FILE REFERENCE: 10001 NP
CURRENT APPLICATION NUMBER: US/11/169,041
PENDING FILING DATE: 2005-06-28
PRIOR APPLICATION NUMBER: 60/584,405
PRIOR FILING DATE: 2004-06-30
NUMBER OF SEQ ID NOS: 527
SOFTWARE: PatentIn version 3.2
SEQ ID NO 162
LENGTH: 543
TYPE: PRT
ORGANISM: Homo sapiens
US-11-169-041-162

Query Match 45.3%; Score 34; DB 7; Length 543;

Best Local Similarity 46.2%; Pred. No. 1.1e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 GGRKMLKSTRROR 15
DB 36 GGRILRORRROLR 48

Search completed: February 4, 2006, 03:47:59
Job time : 11.5556 secs

THIS PAGE BLANK (USPTO)